

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

**VÝZKUM MODERNÍCH ARTIKULAČNÍCH PARAMETRŮ
PRO ANALÝZU HYPOKINETICKÉ DYSARTRIE**

RESEARCH OF MODERN ARTICULATION FEATURES FOR THE ANALYSIS OF HYPOKINETIC
DYSARTHRIA

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Filip Vrba

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Zoltán Galáž

BRNO 2019

Bakalářská práce

bakalářský studijní obor **Audio inženýrství**
Ústav telekomunikací

Student: Filip Vrba

ID: 186641

Ročník: 3

Akademický rok: 2018/19

NÁZEV TÉMATU:

Výzkum moderních artikulačních parametrů pro analýzu hypokinetické dysartrie

POKYNY PRO VYPRACOVÁNÍ:

V rámci bakalářské práce budou navrženy a implementovány nové, moderní řečové parametry kvantifikující poruchy artikulace u pacientů s hypokinetickou dysartrií. Tyto parametry budou hodnotit: 1) hybnost mluvidel, 2) tremor úst a brady a 3) tempo řeči. Dále budou navrženy zcela nové tzv. globální parametry kombinující tyto tři vzpomínané aspekty za účelem robustního popisu artikulace při hypokinetické dysartrii. Tyto parametry budou porovnány s tzv. konvenčními parametry za účelem identifikace Parkinsonovy nemoci.

DOPORUČENÁ LITERATURA:

[1] PELED, Abraham. Digital signal processing: theory, design, and implementation. New York: John Wiley & Sons, 1976.

[2] PSUTKA, Josef. Mluvíme s počítačem česky. Praha: Academia, 2006. Česká matice technická, roč. 111, č. spisu 502. ISBN 80-200-1309-1.

Termín zadání: 1.2.2019

Termín odevzdání: 27.5.2019

Vedoucí práce: Ing. Zoltán Galáž

Konzultant:

prof. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce se zabývá hypokinetickou dysartrií, jakožto poruchou motorické realizace řeči, která se vyskytuje u přibližně 70% pacientů s Parkinsonovou nemocí (PN). V práci je navržena a analyzována dvojice nově vytvořených řečových parametrů pro kvantifikaci artikulace při HD. Tyto parametry byly validovány na nahrávkách zdravých řečníků a pacientů s PN. V teoretické části práce jsou popsány konvenční a využití metody zpracování řečových signálů, parametrizace a statistické analýzy. V části realizace systému je popsán praktický návrh nových parametrů a také použité metody jejich statistického zhodnocení pomocí korelační analýzy a strojového učení. Cílem práce je návrh nových řečových parametrů pro diagnostiku HD. Navržený systém byl realizován v programovém prostředí MATLAB.

KLÍČOVÁ SLOVA

Hypokinetická dysartrie, Parkinsonova nemoc, parametrizace řeči, objektivní analýza, signálové zpracování řeči, statistická analýza, strojové učení

ABSTRACT

This thesis deals with hypokinetic dysarthria, as a disorder of motor speech, which occurs in approximately 70 % of patients with Parkinson's disease (PD). Two newly designed speech parameters for quantification of articulation within HD are analysed in this thesis. These parameters were validated on recording of both healthy and PD speakers. The theoretical part describes conventional and used methods of speech signal processing, parameterization and statistical analysis. In the part of the system implementation is described practical design of new parameters and also methods of their statistical evaluation by correlation analysis and machine learning. The aim of this work is to design new speech parameters for HD diagnostics. The proposed system was implemented in MATLAB software environment.

KEYWORDS

Hypokinetic dysarthria, Parkinson's disease, speech parameterization, objective analysis, speech signal processing, statistical analysis, machine learning

VRBA, Filip. *Výzkum moderních artikulačních parametrů pro analýzu hypokinetické dysartrie*. Brno, 2019, 65 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Zoltán Galáž

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Výzkum moderních artikulačních parametrů pro analýzu hypokinetické dysartrie“ jsem vypracoval(a) samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autora(-ky)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Zoltánu Galážovi za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno

.....

podpis autora(-ky)

PODĚKOVÁNÍ

Výzkum popsany v této bakalářské práci byl realizován v laboratořích podpořených z projektu SIX; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Brno

.....
podpis autora(-ky)

OBSAH

Úvod	11
1 Parkinsonova nemoc	13
1.1 Parkinsonova nemoc a její příčiny	13
1.2 Příznaky Parkinsonovy nemoci	13
2 Hypokinetická dysartrie	14
2.1 Oblast fonace	14
2.2 Oblast artikulace	14
2.3 Oblast prozodie	15
3 Klinický popis pacientů	16
3.1 Test 3F: Dysartrický profil	16
3.2 Databáze řečových nahrávek	17
4 Konvenční postupy kvantifikace	
dysartické artikulace	18
4.1 Předzpracování řečových signálů	18
4.1.1 Odstranění stejnosměrné složky	18
4.1.2 Preemfázová filtrace	18
4.1.3 Segmentace řeči	19
4.2 Parametrizace řečových signálů	21
4.2.1 Lineární predikční analýza LPC	21
4.2.2 Formantové kmitočty	21
4.2.3 Parametr VSA	23
4.2.4 Parametr FCR	24
4.2.5 Parametr VAI	24
5 Nové postupy kvantifikace	
dysartrické artikulace	25
5.1 Parametr Formant Variability	25
5.2 Parametr Speech Articulation	26
6 Realizace systému	27
6.1 Blokové schéma systému	27
6.2 Popis systému	28
6.2.1 Načtení a zpracování dat	28
6.2.2 Předzpracování	28

6.2.3	Výpočet akustických parametrů	28
6.2.4	Post-processing	29
6.3	Statistická analýza	29
6.3.1	Porovnání akustických parametrů	29
6.3.2	Korelační analýza	30
6.3.3	Parsonův korelační koeficient	30
6.3.4	Spearmanův korelační koeficient	30
6.3.5	Výpočet korelace parametrů	31
6.3.6	Strojové učení	31
7	Výsledky	35
7.1	Korelační analýza	35
7.2	Klasifikační analýza	37
7.3	Regresní analýza	38
8	Závěr	44
	Literatura	46
	Seznam symbolů, veličin a zkratk	49
	Seznam příloh	50
A	Fonace českých samohlásek	51
B	Formanty českých samohlásek	52
C	Zdrojové kódy parametrů	54
C.1	Kód parametru funkce Formant Variability	54
C.2	Kód parametru funkce Speech Articulation	55
D	Porovnání parametrů	58
D.1	Formant variability	58
D.2	Speech articulation	63
E	Obsah přiloženého CD	65

SEZNAM OBRÁZKŮ

4.1	Modulová charakteristika preem. filtru	19
4.2	Modulová charakteristika váhovacích posloupností	20
4.3	Modulové spektrum samohlásky [a	22
4.4	Pětiúhelník parametru VSA	24
6.1	Blokové schéma systému	27
6.2	Tabulka binární klasifikace	32
A.1	Časový průběh českých samohlásek	51
B.1	Modulové spektrum české samohlásky a	52
B.2	Modulové spektrum české samohlásky e	52
B.3	Modulové spektrum české samohlásky i	53
B.4	Modulové spektrum české samohlásky o	53
B.5	Modulové spektrum české samohlásky u	53
D.1	Sklon kumulativních funkcí a počet špičkových hodnot pro F_1, F_2, F_3	58
D.2	Minima a maxima pro F_1, F_2, F_3	59
D.3	Průměr a směrodatná odchylka pro F_1, F_2, F_3	60
D.4	Rozsah pro F_1, F_2, F_3 a hodnoty kvartilů Q_1, Q_2 a Q_3 pro F_1	61
D.5	Hodnoty kvartilů Q_1, Q_2 a Q_3 pro F_2 a F_3	62
D.6	Sklon kumulativní funkce, špičkové hodnoty, minima, maxima, průměr a směrodatná odchylka pro F	63
D.7	Rozsah a hodnoty Q_1, Q_2 a Q_3 pro F	64

SEZNAM TABULEK

3.1	Vybrané části protokolu řečových cvičení.	16
3.2	Statistické charakteristiky PARCZ databáze.	17
4.1	Hellwagův trojúhelník	22
7.1	Korelační analýza parametru Formant Variability.	36
7.2	Korelační analýza parametru Speech Articulation.	36
7.3	Klasifikační analýza FV pro jednu proměnnou	37
7.4	Klasifikační analýza SA pro jednu proměnnou	37
7.5	Klasifikační analýza FV se všemi parametry	38
7.6	Klasifikační analýza SA se všemi parametry	38
7.7	Regresní analýza FV s jedním parametrem pro UPDRSIII	39
7.8	Regresní analýza SA jedním parametrem pro UPDRSIII	39
7.9	Regresní analýza FV s jedním parametrem pro UPDRSIV	39
7.10	Regresní analýza SA s jedním parametrem pro UPDRSIV	40
7.11	Regresní analýza FV s jedním parametrem pro FOG	40
7.12	Regresní analýza SA s jedním parametrem pro FOG	40
7.13	Regresní analýza FV s jedním parametrem pro DX	41
7.14	Regresní analýza SA s jedním parametrem pro DX	41
7.15	Regresní analýza FV s jedním parametrem pro ACE-R	42
7.16	Regresní analýza SA s jedním parametrem pro ACE-R	42
7.17	Regresní analýza FV se všemi parametry pro dané škály	43
7.18	Regresní analýza SA se všemi parametry pro dané škály	43

ÚVOD

Řeč je bezpochyby jeden z nejdůležitějších nástrojů lidské sociální interakce. Je to po tisíciletí budovaný a pilovaný dorozumívací prostředek, jenž má v naší civilizaci zastoupení v nesčetných formách. Tento jemný a velmi komplikovaný systém, pomocí kterého by se člověk jak ho známe již neobešel, vychází z velké řady mnohdy na sobě závislých dílčích procesů a úkonů. Hlasivky, hrtan, jazyk, plíce, dutina ústní a také mozek, to vše jsou kooperující součásti, které dávají za vznik řeči. V případě nedokonalé funkčnosti jedné či více z nich, dochází v promluvě k nežádoucím jevům a mluvíme tak o poruše řeči.

Mezi nemoci které s poruchami řeči souvisí, patří také Parkinsonova nemoc. Toto neurodegenerativní onemocnění má mimo další fyziologické projevy za následek poruchu motorické realizace řeči označovanou jako hypokinetická dysartrie. Právě touto nemocí se tato práce zabývá. Vlivem zhoršené schopnosti kontroly svalů dochází k patologickým projevům v oblasti fonace, prozodie, plynulosti řeči, ale také artikulace na kterou se tato práce blíže zaměřuje. Pomocí analýzy řeči v těchto oblastech je možné vysledovat patologické jevy, které mohou napomoci k diagnostice, hodnocení a také monitorování tohoto onemocnění. Pro potřeby analýzy řeči existují tzv. parametry, které se svou komplexností pohybují od sledování běžných příznaků, až po velmi složité metody které tyto příznaky kombinují.

V této práci je zpočátku věnována pozornost Parkinsonově nemoci a jejím příznakům a projevům. Dále je zde popsána hypokinetická dysartrie a její vliv na artikulační, fonační a prozodickou oblast řeči. Následující část se věnuje dysartrickému profilu a tzv. 3F testu používanému pro jeho kvantifikaci. Také je zde představena řečová databáze PARCZ (Czech Parkinsonian Speech Database), která slouží jako podklad pro analýzu patologické řeči. Ve čtvrté kapitole se práce zaměřuje na konvenční postupy kvantifikace dysartrické artikulace, jako jsou předzpracování řečových signálů: odstranění stejnosměrné složky, segmentace, nebo preemfázová filtrace. V následující části je již nastíněna samotná parametrizace řečových signálů. Kromě popisu důležitých nástrojů jako lineární predikční analýza, či obeznámení se s vokální oblastí hlasového traktu, jsou zde také popsány některé konvenčně využívané parametry založené na popisu rezonancí v dutině hrdelní, ústní a nosní, známé jako tzv. formanty. V páté kapitole je popsán návrh matematické realizace dvojice nových parametrů, které jsou určeny pro kvantifikaci dysartrické řeči se zaměřením na oblas artikulace. Následující část se pak věnuje popisu realizace těchto parametrů v programovém prostředí Matlab. Pozornost je zde věnována popisu jednotlivých bloků navrženého systému jako: načítání dat, předzpracování signálu, výpočtu akustických parametrů a následnému post-processingu. Dále jsou zde popsány metody testování navržených parametrů pomocí metod korelační analýzy a strojového učení

(SU). V rámci strojového učení jsou zde popsány konkrétní využití metody jako je klasifikační a regresní analýza. Také je zde uvedena a popsána metoda srovnání akustických parametrů pomocí krabicových diagramů. Poslední kapitola se věnuje výsledkům daných analýz.

1 PARKINSONOVA NEMOC

1.1 Parkinsonova nemoc a její příčiny

Parkinsonova nemoc (Morbus Parkinsoni) je neurodegenerativní onemocnění, které postihuje centrální nervovou soustavu člověka [4]. Poprvé bylo popsáno v práci „An Essay on the Shaking Palsy“ Jamesem Parkinsonem v roce 1817. Nemoc se vyznačuje především progresivním úbytkem dopaminergických neuronů, zejména v oblasti tzv. *substantia nigra pars compacta*. Dále také dochází k dalším progresivním deficitům v několika neurotransmiterových systémech. Porucha dopaminergní inervace striata, vede k poruše funkce zpětnovazebních okruhů zvaných extrapyramidový systém.

Úbytek dopaminu vede ke většině parkinsonských symptomů a je zároveň podkladem léčebného účinku dopaminergní medikace na motorické symptomy PN [13]. Nedostatek dopaminu, který přirozeně zajišťuje přenos signálů mezi neurony zapříčiňuje různé motorické symptomy. Ty způsobují postupnou ztrátu schopnosti ovládnutí i kontroly pohybu pacienta. Bylo prokázáno, že progres PN je nelineární. Na počátku onemocnění můžeme pozorovat rychlý nástup motorických symptomů, ale v dalších stádiích nejsou změny natolik prudké a výskyt nových příznaků a jejich výrazný vývoj obvykle klesá [11].

1.2 Příznaky Parkinsonovy nemoci

Příznaky PN jsou u každého člověka individuální a projevují se odlišně. Stav pacientů se v průběhu dne výrazně střídá, dojem téměř zdravého člověka tak u pacienta během několika hodin naopak střídají silné projevy symptomů.

Nejčastějšími projevy jsou třes (tremor), pocity únavy, ztuhlost (rigidita) v končetinách a zádech. Dale pak křeče, zácpa, problémy se spánkem a poruchy rovnováhy. Rigidita se obvykle poprvé projevuje na jedné z končin či jedné straně těla, někdy také na končetinách do kříže a postupně přechází do dalších oblastí. Zvýšené napětí svalů způsobené ztuhlostí zapříčiňuje bolesti a často také neschopnost dovést pohyb do krajní polohy. V některých případech může také dojít ke kompletní ztuhlosti dočasně znemožňující další pohyb.

V symptomech následuje zpomalení chůze, zhoršená mimika, inkontinence. Jedním z prvních varovných příznaků je klidový třes rukou, či také zhoršení čichu a chuti. Parkinsonický třes může zachvátit taky jazyk, rty a krk. Tím v kombinaci s rigiditou způsobuje značné problémy s komunikací. Pacient není schopen správně artikulovat a výsledná řeč se jeví jako tichá, sekaná či monotónní.

2 HYPOKINETICKÁ DYSARTRIE

Přibližně 70 % pacientů postižených Parkinsonovou nemocí má problémy s motorickými aspekty řeči [8], které lze shrnout pod označením *hypokinetická dysartrie* (HD). Pojem hypokinetická dysartrie lze definovat jako souhrnné označení pro neurogenní poškození centrálního či periferního nervového systému, způsobující poruchu realizace řeči a zapříčiňující celkovou neschopnost kontroly řečových mechanismů.

Z pohledu klinické praxe je možné rozsah poruchy rozdělit na lehký, střední a těžký stupeň závažnosti poškození řeči. Nejtěžší případ při kterém dochází k úplné neschopnosti řeči se označuje jako tzv. *anartrie* [11]. Hypokinetická dysartrie se projevuje v oblastech fonace, artikulace, prozodie, plynulosti řeči a mimice obličeje neboli tzv. faciokinezi.

2.1 Oblast fonace

Fonace zahrnuje proces vytváření zvuku v hlasivkách, případně v celém hrtanu, který je následně dále modulován v dutině ústní, hrdelní a nosní. V příloze ?? je možné vidět příklady časového průběhu a také jednostranného modulového spektra fonace všech pěti českých samohlásek: [a], [e], [i], [o], [u].

Fonace je obvykle první oblastí, ve které se symptomy HD projevují. Vlivem rigidity svalů dochází k nesprávnému fungování hlasivek, především k neschopnosti jejich plného uzavření. Turbulentní proudění vzduchu z plic kolísá a projevuje se zde tzv. hlasový třes [7]. Dále můžeme u pacientů pozorovat zvýšenou hodnotu střední frekvence základního tónu F_0 a to jak při delších větných úsecích, tak při jednotlivých prodloužených vokálech. Bylo prokázáno, že s prodlužující se dobou onemocnění tato hodnota narůstá [16]. Z toho lze usuzovat, že je hodnota F_0 vhodným parametrem pro určení pokročilosti onemocnění.

U pacientů s PN se projevují tyto vady řeči související s fonací: *hypofonie* neboli snížená intenzita hlasu, *dysfonie* neboli degradace kvality hlasu (chraptivost, drsnost, atd.), *hypernasalita* neboli prudké uvolnění dechového sloupce skrze dutinu nosní, atd. Tyto průvodní jevy v značné míře znehodnocují kvalitu lidské promluvy a mohou snížit srozumitelnost a také přirozenost řeči.

2.2 Oblast artikulace

Oblast artikulace je druhou nejvýznamnější oblastí ve které se symptomy Parkinsonovy nemoci a s ní spojené hypokinetické dysartrie projevují. Porucha se zde vyskytuje především při vyslovování souhlásek, neboli *kosonant* [25]. Problémy se

vyskytují nejvýrazněji u vyslovování ploziv [p], [t], [b], [d], [g] [3], zvláště u plozivy [k]. Tyto problémy způsobují řeč působící nepřirozeně a nesrozumitelně. Navině je nesprávná práce aktivních artikulačních orgánů: jazyka, rtů, hrtanové příklopky a hrtanu.

Pro hodnocení kvality artikulace se používají zejména tzv. diadochokinetické úlohy neboli zkráceně DDK (*Diadochokinetic Tasks*). Pomocí repetitivních frází střídající kombinaci slabik s plozivou a vokálem jako např. „pa–ta–ka–pa–ta–ka–pa–ta–ka...“ [5] a jiné. Pacienti postižení HD nejsou při těchto úlohách schopni správně provádět výrazné změny poloh artikulačních orgánů. Stává se také, že pacienti udržují kadenci výslovnosti na úkor srozumitelnosti.

Pro srozumitelnou řeč je zapotřebí zapojení rtů, jazyka, měkkého patra a veškerých prostorů umožňujících rezonanci; ústní dutina, hltan a v menší míře hrudní dutina. Správné vyslovování vokálů se odvíjí právě od objemu hrtanu a dutiny ústní, kde dochází rezonancím které můžeme měřit pomocí tzv. formantových kmitočtů. Ty jsou prostorem těchto dutin zásadně formovány [20, 21, 6]. V příloze B je možné vidět příklady odhadu prvních dvou formantových kmitočtů pro všech pět českých samohlásek: [a], [e], [i], [o], [u].

2.3 Oblast prozodie

Poslední ze tří oblastí projevů HD je oblast prozodie řeči. Mezi prozodické faktory řeči se řadí přízvuk, melodie, intonace, pauzování a také tempo řeči. Tato oblast úzce souvisí s oblastí fonace, jelikož mezi její základní tzv. *suprasegmentální rysy* patří kmitočet základního tónu, intenzita a tempo řeči. Důležitým parametrem je mimo jiné i její rytmus. Dále se u pacientů projevuje zpomalování řeči, neboli tzv. *bradyfemie*, zrychlování řeči neboli tzv. *tachyfemie*, zrychlené opakování stejných slabik na začátku, uprostřed nebo na konci slov neboli tzv. *palilalie* a také neplánované či nezáměrné pauzy v projevu neboli tzv. *hezitace*.

Při narušení prozodických faktorů řeči dochází k výrazné změně emočního ladění a řeč je tak ochuzena o zdůraznění či citové zabarvení promluvy. Takto postižení pacienti jej proto nemohou adekvátně vyjádřit [2] a jejich řeč se stává méně srozumitelnou a přirozenou. Mluvíme tak o tzv. *dysprozodii* [16]. Dysprozodie se tedy vyznačuje monotónností řeči (jak ve výšce hlasu, tak v intenzitě) a také nepřirozeným tempem a pauzováním.

3 KLINICKÝ POPIS PACIENTŮ

3.1 Test 3F: Dysartrický profil

Hypokinetická dysartrie je komplexní poškození řeči, které může ovlivnit kterýkoliv systém podílející se na tvorbě řeči (respirace, fonace, artikulace, prozodie). Z tohoto důvodu je kvantifikace tohoto onemocnění poměrně náročná a vyžaduje komplexní sadu řečových cvičení, které jsou schopny jednotlivé příznaky této nemoci zdůraznit. V České Republice je pro tyto účely běžně používán tzv. 3F test: Dysartrický profil.

V této práci je využit protokol řečových cvičení navržen Klinikou logopedkou PhDr. Mgr. Milenou Košťálovou z Neurologické kliniky Lékařské fakulty Masarykovy univerzity. Tento protokol obsahuje celkem 91 řečových úloh. Protokol byl navržen tak, aby bylo možné odhadnout jednotlivé výsledky 3F testu a také výsledný dysartrický index. Vybrané části protokolu jsou uvedeny v tab. 3.1. V práci jsou uvedena pouze řečová cvičení, která jsou následně použita pro výpočet řečových parametrů.

Tab. 3.1: Vybrané části protokolu řečových cvičení.

Kód	Úloha
Diadochokineze s fonací	
7.4-1	Pacient na jeden nádech a po co nejdelší dobu rychle opakuje „ba-ba-ba“
7.4-2	Pacient na jeden nádech a po co nejdelší dobu rychle opakuje „p-t-k“
7.4-3	Pacient na jeden nádech a po co nejdelší dobu rychle opakuje „pa-ta-ka“
Srozumitelnost předříkávaných vět	
9.2-1	Pacient opakuje podle vyšetřujícího větu „ <i>Do čtvrt hodiny tam byla smršť.</i> “ vyslovenou jeho normálním způsobem mluvy.
9.2-2	Pacient opakuje podle vyšetřujícího větu „ <i>Prohovořte to s ním dopodrobna.</i> “ vyslovenou jeho normálním způsobem mluvy.
9.2-3	Pacient opakuje podle vyšetřujícího větu „ <i>Při ústupu pluku duní bubny.</i> “ vyslovenou jeho normálním způsobem mluvy.
9.2-4	Pacient opakuje podle vyšetřujícího větu „ <i>Kuchařští učni nejsou jak zlatničtí.</i> “ vyslovenou jeho normálním způsobem mluvy.
9.2-5	Pacient opakuje podle vyšetřujícího větu „ <i>Celý večer se učí sčítat.</i> “ vyslovenou jeho normálním způsobem mluvy.

Hypokinetická dysartrie je pomocí 3F testu hodnocena pomocí 3-stupňové škály u celkem 30 položek: faciokineze, fonorespirace (hodnotí stav artikulačních orgánů jako jsou rty, jazyk nebo měkké patro) a fonetika (přesnost opakování souhlásek/samohlásek, artikulace při čtení textu...), srozumitelnost (četba/předříkávání slov nebo vět, spontánní řeč, atd.) a prozodie řeči (intonace, rytmus, pauzování...). Výsledkem 3F testu je *dysartrický profil*, který udává tzv. *dysartrický index DX*. Tento index nabývá hodnot v intervalu $\langle 0;90 \rangle$:

- 85–90 *Bez poruchy.*
- 75–84 *Velmi lehká dysartrie.*
- 57–74 *Lehká dysartrie.*
- 36–56 *Středně těžká dysartrie.*
- 17–35 *Těžká dysartrie.*
- 0–16 *Velmi těžká dysartrie (anartrie).*

3.2 Databáze řečových nahrávek

V této práci bude využito databáze řečových nahrávek, které byly pořízeny na pracovišti I. Neurologické kliniky Lékařské fakulty Masarykovy univerzity a Fakultní nemocnice u sv. Anny v Brně. Tato databáze se nazývá PARCZ (Czech Parkinsonian Speech Database) a obsahuje dva typy řečníků: pacienti s Parkinsonovou nemocí (100 řečníků: 60 mužů, 40 žen) a věkově vázané kontrolní řečníky (52 řečníků: 26 mužů, 26 žen). Statistické charakteristiky této databáze je možné vidět v tab. 3.2.

Tab. 3.2: Statistické charakteristiky PARCZ databáze.

	Počet		Průměrný věk		Věkový rozsah		STD věku	
	Muži	Ženy	Muži	Ženy	Muži	Ženy	Muži	Ženy
Zdraví	26	26	65,58	61,81	49–83	45–87	8,90	9,05
Nemocní	60	40	66,28	69,98	46–87	49–86	8,63	7,74

4 KONVENČNÍ POSTUPY KVANTIFIKACE DYSARTICKÉ ARTIKULACE

4.1 Předzpracování řečových signálů

Nezbytnou součástí zpracování řečových signálů, je před jejich použitím řádná úprava. Číslíkově zpracovávaný signál je potřeba zbavit negativních vlivů způsobených nahráváním (okolní hluk, šum, zkreslení při záznamu, vlastnosti zesilovače či mikrofону atd.). Základem pro kvalitní řečový záznam je odpovídající záznamový aparát (zesilovač, mikrofon) a také vhodně zvolené prostředí - nejlépe bezodrazová komora [18, 24].

4.1.1 Odstranění stejnosměrné složky

Při parametrizaci řečového signálu je běžnou praxí tzv. odstranění stejnosměrné složky signálu. Tato složka je totiž pro výpočet většiny parametrů nepodstatná a v některých případech může být nežádoucí. Odstranění stejnosměrné složky neboli tzv. *ustředění* se provádí pomocí následujícího vztahu:

$$s^*[n] = s[n] - \frac{1}{N} \sum_{n=0}^{N-1} s[n], \quad (4.1)$$

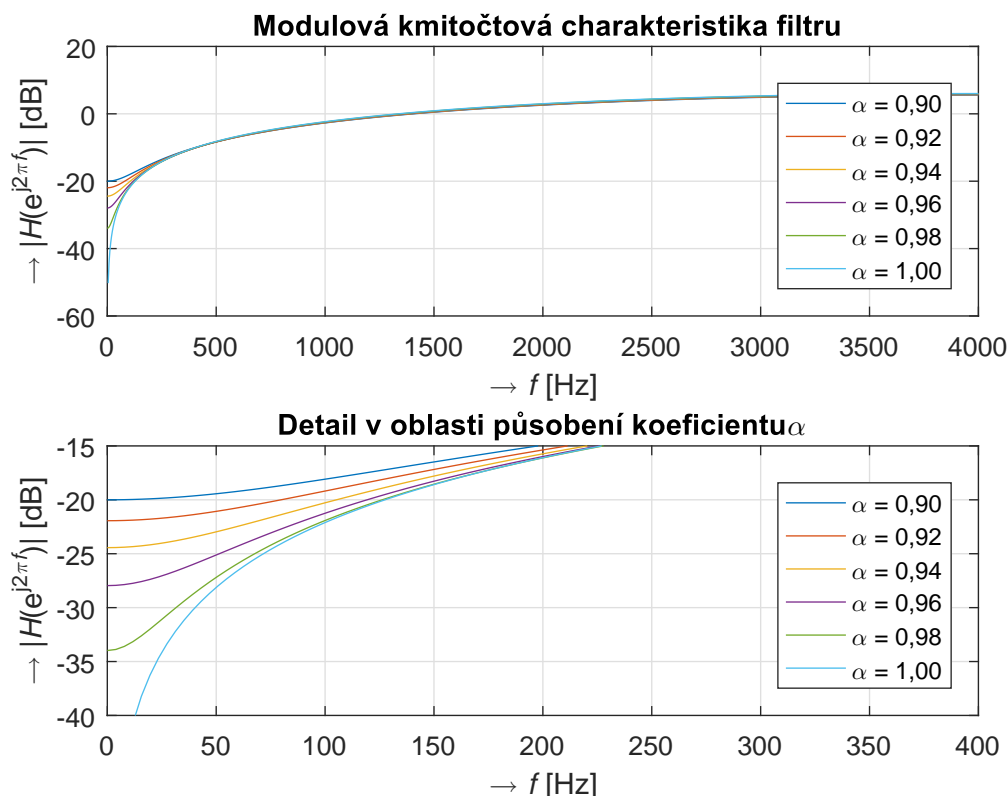
kde N je počet vzorků signálu, $s[n]$ jsou jednotlivé vzorky vstupního signálu a $s^*[n]$ jsou jednotlivé vzorky výstupního signálu.

4.1.2 Preemfázová filtrace

Preemfázová filtrace se při parametrizaci řeči využívá ke kompenzaci kmitočtové charakteristiky prostředí. Přenosové prostředí má charakter dolní propusti. Utlučuje tím pádem vyšší kmitočty, které jsou ale pro správnou srozumitelnost řeči stěžejní. Preemfázový filtr má tedy charakter kompenzační horní propusti (vyrovňuje spektrum řečového signálu potlačené v důsledku šíření signálu). Přenosová funkce klasické implementace preemfázového filtru je popsána jako systém prvního řádu [24]:

$$H(z) = 1 - \alpha z^{-1}, \quad (4.2)$$

kde α je tzv. koeficient preemfázového filtru a běžně nabývá hodnot 0.9 až 1. Příklad modulové kmitočtové charakteristiky preemfázového filtru pro různé hodnoty koeficientu $\alpha = \{0.90, 0.92, 0.94, 0.96, 0.98, 1.00\}$ je možné vidět na Obr. 4.1.



Obr. 4.1: Modulová kmitočtová charakteristika preemfázového číslicového filtru, parametry: $Fs = 8000$ Hz, $\alpha = \{0.90, 0.92, 0.94, 0.96, 0.98, 1.00\}$.

4.1.3 Segmentace řeči

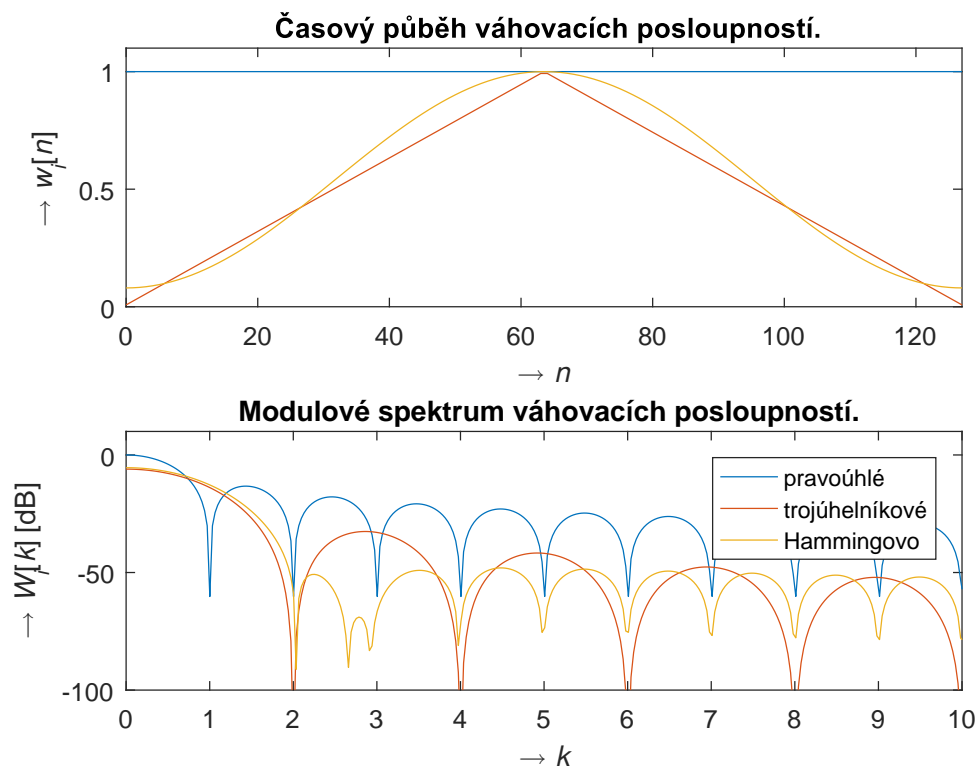
Signál je také před dalším zpracováním potřeba rozdělit na segmenty, neboli rámce. Řečový signál považujeme za náhodný, je tedy pro konvenční metody odhadu řečových parametrů jen těžce použitelný. Proto je nutné signál upravit, aby jej bylo možné považovat za stacionární a tedy použitelný pro proces řečové parametrizace. Tohoto stavu docílíme díky rozdělení signálu na kratší úseky tak, že jeho statistické charakteristiky nejsou závislé na posunutí počátku časové osy.

Délka daných rámců by měla být dostatečně malá, aby byl signál v daném úseku stacionární, zároveň však dostatečně velká pro odečítání charakteristik řeči potřebných požadovanými parametry. Ideální je délka odpovídající setrvačnosti hlasového ústrojí pohybující se kolem 20 – 25 ms. Pro vytvoření rámce využíváme tzv. váhovací posloupnosti (váhovací okna) a to především: tzv. pravoúhlé, trojúhelníkové, Blackmanovo, Hannovo, Hammingovo, atd. Pro zpracování dysartrické řeči je nej-

používanější okno Hammingovo, které je definováno jako [18]:

$$W[k] = \begin{cases} 0.54 - 0.46 \cos \frac{2\pi n}{N}, & \text{pro } n=1,2,\dots,N \\ 0, & \text{jinde.} \end{cases} \quad (4.3)$$

Zde je k tzv. index vzorku. Na Obr. 4.2 je zobrazen časový průběh a modulová kmitočtová charakteristika několika váhovacích posloupností. Konkrétně se jedná o okna: pravoúhlé, trojúhelníkové a Hammingovo. Jak je možné vidět, Hammingovo okno poskytuje nejlepší potlačení signálu v tzv. vedlejších lalocích, čímž zabráňuje tzv. prosakování spektrálních složek z vedlejších segmentů. V oblasti zpracování dysartrické řeči se běžně používá délka okna 20 – 25 jak již bylo zmíněno výše a překrytí 50 %.



Obr. 4.2: Časový průběh a modulová kmitočtová charakteristika váhovacích posloupností. Parametry: $F_s = 8000$ Hz, $NFFT = 2048$, k je index vzorků posloupnosti.

4.2 Parametrizace řečových signálů

4.2.1 Lineární predikční analýza LPC

Lineární predikce, neboli *Linear Predictive Coding* (LPC), je algoritmus ztrátové komprese na autoregresivním modelu (*Auto-Regressive AR*). Algoritmus během tzv. dopředné a zpětné predikce využívá dvojici číslicových filtrů typu FIR (*Finite Impulse Response*), neboli filtry s konečnou impulzní odezvou. Lineární predikční analýza odhaduje aktuální vzorek řeči $s[n]$ z M předešlých vzorků. Pro dopřednou lineární predikční analýzu platí vztah [24]:

$$\tilde{s}[n] = -(a_1 s[n-1] + \dots + a_M s[n-M]) = -\sum_{i=1}^M a_i s[n-i], \quad (4.4)$$

kde $\tilde{s}[n]$ je predikovaný vzorek, M udává počet předchozích vzorků a a_i koeficienty filtru (lineární predikční koeficienty). Základní blok parametrického modelu pro syntézu řeči, který využívá lineární predikční koeficienty a model hlasového traktu 10. řádu má tvar přenosové funkce [24]:

$$H(z) = \frac{c_0}{d_0 + d_1 z^{-1} + \dots + d_{10} z^{-10}} = \frac{G}{1 + \sum_{i=1}^{10} a_p[i] z^{-i}} = \frac{G}{A_p(z)} = \frac{Y(z)}{X(z)}. \quad (4.5)$$

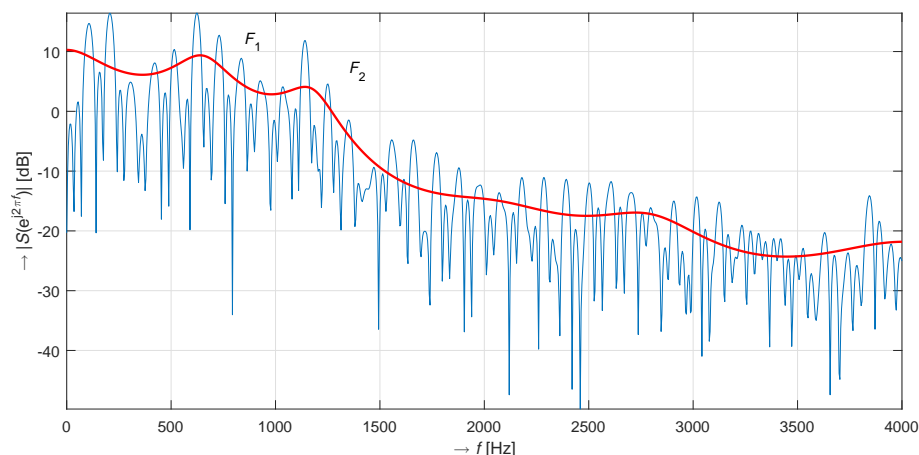
Koeficienty $a_p[i]$ jsou lineární dopředné predikční koeficienty. Tento LPC model aproximuje formanty, tj. rezonanční vrcholy, které se ve spektru řečového signálu objevují [24]. Pomocí tohoto algoritmu je možné poměrně spolehlivě odhadovat parametry řeči bez velkých nároků na operační paměť.

4.2.2 Formantové kmitočty

Při tvorbě lidského hlasu je proud vzduchu z plic hnán přes hlasivky do nadhrtanových dutin (hrdelní, nosní a ústní). Signál vycházející z hlasivek se označuje jako základní hlasový tón. Při jeho průchodu nadhrtanovými dutinami v nich dochází k rezonancím na kmitočtech, které se nazývají formanty (F_1, F_2, \dots, F_n) [18]. Hodnoty formantů a odpovídající prostor v dutinách ovlivňují svou polohou mluvidla. Díky tomu je pomocí formantu např. možné sledovat hybnost jazyka. Platí následující přiřazení formantů a dutin [17]:

1. F_1 – neboli frekvence prvního formantu (dutina hrdelní)
2. F_2 – neboli frekvence druhého formantu (dutina ústní)
3. F_3 – neboli frekvence třetího formantu (dutina nosní)

Při parametrizaci řečového signálu se formanty obvykle doplňují také o odpovídající šířky pásem B_1, B_2, B_3 . Nevýhodou využití formantů ve zpracování řeči je jejich výrazná změna v závislosti na věku a pohlaví zkoumané osoby [18, 15]. Bylo



Obr. 4.3:]

Modulové spektrum české samohlásky [a] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC Parametry: $Fs = 8000$ Hz

Tab. 4.1: Hellwagův trojúhelník (klasifikace vokálů z hlediska polohy jazyka). Parametry: F_1 = kmitočet prvního formantu, F_2 = kmitočet druhého formantu.

		$F_2 \rightarrow$			
		Zadní		Střední	Přední
$F_1 \uparrow$	Nízké	[a]			
	Středové	[o]		[e]	
	Vysoké	[u]			[i]

zjištěno, že například první dva formanty F_1 a F_2 jsou nejvíce zodpovědné za tvorbu vokálů, neboli samohlásek [10].

Na obr. 4.3 můžeme vidět příklad modulového spektra české samohlásky [a] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC. Modulová spektra pro všechny české samohlásky se nachází v příloze B. Tyto formanty byly odhadovány programem Praat [1].

Vokály (samohlásky) jsou znělé a vznikají v ustálené poloze hlasového traktu. Jsou buzeny periodickým signálem, který má relativně vysokou energii. V případě horizontálního směru pohybu jazyka rozlišujeme samohlásky přední, střední a zadní. Ve vertikálním směru dělíme na vysoké, středové a nízké. Tuto klasifikaci poprvé popsal Ch. F. Hellwag a závislost se zobrazuje ve formě vokalického trojúhelníku, viz tabulka 4.1. Bylo prokázáno, že horizontální pohyb ovlivňuje velikost druhého formantu a ve vertikálním směru pak formant první [9].

4.2.3 Parametr VSA

Pomocí formantových kmitočtů (F_1 , F_2 a F_3) je možné vyjádřit míru hybnosti jazyka. Na tomto principu je založen příznak VSA, neboli *Vowel Space Area*. Samohlásky (vokály) [a], [e], [i], [o], [u] jsou pro tento účel extrahovány z řeči a poté se odhadnou hodnoty prvních dvou formantů (F_1 , F_2). Hodnoty se zanesou do grafu, což vytvoří pětiúhelník, viz Obr. 4.4, jehož obsah odpovídá parametru VSA. Výpočet parametru VSA je popsán následujícími rovnicemi [14]:

$$|v_{xy}| = \sqrt{(F_{1x} - F_{1y})^2 + (F_{2x} - F_{2y})^2} \quad (4.6)$$

$$a_1 = |v_{ue}|, b_1 = |v_{ei}|, c_1 = |v_{iu}| \quad (4.7)$$

$$a_2 = |v_{uo}|, b_2 = |v_{oe}|, c_2 = |v_{eu}| \quad (4.8)$$

$$a_3 = |v_{oa}|, b_3 = |v_{ae}|, c_3 = |v_{eo}| \quad (4.9)$$

$$s_i = \frac{a_i + b_i + c_i}{2} \quad (4.10)$$

$$Tr_i = \sqrt{s_i(s_i - a_i)(s_i - b_i)(s_i - c_i)} \quad (4.11)$$

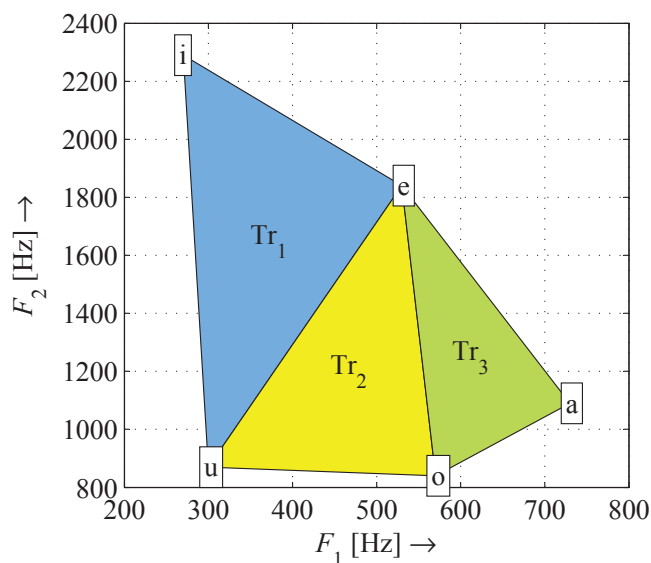
$$VSA = \sum_{i=1}^3 Tr_i \quad (4.12)$$

kde F_{ix} a F_{iy} vyjadřují frekvence i-tého formantu odpovídající samohlásky. Tr_i je obsah tzv. trojúhelníkového VSA (tVSA) popsaného rovnicí [20]:

$$VSA = ABS((F_{1i}(F_{2a} - F_{2u}) + F_{1a}(F_{2u} - F_{2i}) + F_{1u}(F_{2i} - F_{2a}))/2), \quad (4.13)$$

kde ABS je absolutní hodnota.

Někdy je možné použít také parametr lnVSA, který se počítá obdobně jako parametr VSA, nicméně hodnoty všech formantových kmitočtů se před výpočtem logaritmují. U pacientů postižených hypokinetickou dysartrií se v důsledku poškozené artikulace předpokládá snížená hybnost mluvidel, zejména jazyka. Parametry VSa i lnVSa by tedy měly oproti zdravým řečníkům dosahovat výrazně nižších hodnot [7].



Obr. 4.4: Pětúhelník zobrazující základ pro výpočet artikulačního parametru VSA. Parametry: F_1 = kmitočet prvního formantu, F_2 = kmitočet druhého formantu.

4.2.4 Parametr FCR

Aby se snížila závislost parametru VSA na pohlaví řečníka (tzv. kovariát¹), byl navržen příznak FCR, neboli *Formant Centralization Ratio* [20], který je definován následujícím vztahem:

$$FCR = \frac{F_{2u} + F_{2a} + F_{1i}}{F_{2i} + F_{1a}}. \quad (4.14)$$

Jednotlivé proměnné použité ve vztahu značí následující: F_{2u} – frekvence druhého formantu samohlásky [u], F_{2a} – frekvence druhého formantu samohlásky [a], F_{1i} – frekvence prvního formantu samohlásky [i], F_{2i} – frekvence druhého formantu samohlásky [i] a F_{1a} – frekvence prvního formantu samohlásky [a]. Díky tvaru zlomku je příznak FCR normalizován a nabývá tak na robustnosti [20].

4.2.5 Parametr VAI

VAI, neboli *Vowel Articulation Index* je další parametr nezávislý na pohlaví a je definován jako převrácená hodnota FCR a tudíž popis jednotlivých proměnných je stejný i pro tento parametr.[23]:

$$VAI = \frac{1}{FCR}. \quad (4.15)$$

¹proměnná, která ovlivňuje zkoumanou charakteristiku a je tedy tento vliv zapotřebí uvažovat v samotném výpočtu/analýze

5 NOVÉ POSTUPY KVANTIFIKACE DYSARTRICKÉ ARTIKULACE

V rámci této práce byly navrženy a implementovány dva nové parametry pro kvantifikaci dysartrické řeči se zaměřením na oblast artikulace. Ty se zaměřují na změny a kolísání intenzity prvních tří formantů a v rámci obou parametrů jsou testovány také jejich dílčí akustické parametry. Následující kapitola se věnuje ideovému a předběžnému matematickému návrhu daných parametrů.

5.1 Parametr Formant Variability

Tento nově navržený algoritmus Formant Variability (FV) se zaměřuje na kolísání intenzity signálu prvních tří formantových oblastí lidského hlasu. Je zde využito měření absolutních rozdílů frekvencí F_1 , F_2 a F_3 daných signálů. V rámci parametru je vypočítán vektor absolutních rozdílů po sobě jdoucích formantových frekvencí zredukované medianem těchto absolutních frekvencí. Pro potřebu výpočtu tohoto parametru je potřeba získat hodnoty jednotlivých formantů pro každý sledovaný segment. Dále je zde vytvořeno volitelné nastavení přepočtu signálu na tzv. *semitone*, neboli půltónovou stupnici dle vzorce:

$$F_{semitone}(n) = 12 \cdot \log_2 \left(\frac{F(n)}{F(n-1)} \right). \quad (5.1)$$

Následně je zde využit vzorec k výpočtu absolutních rozdílů těchto hodnot:

$$Diff_{abs}(n) = |F(n) - F(n-1)|, \quad (5.2)$$

kde n reprezentuje n -tý prvek vstupního vektoru hodnot. V parametru jsou poté odstraněny nedůležité hodnoty absolutních rozdílů pomocí odečtení hodnot menších než jejich medián. Parametrem jsou pak kromě samotných absolutních rozdílů hodnot získány také: počet špičkových hodnot signálu tzv. *numpeaks* (počet hodnot ve vektoru již zredukovaných absolutních rozdílů větších, než 2 okolní hodnoty) a také sklon kumulativní funkce (součtu) hodnot po sobě jdoucích absolutních rozdílů. Nižší počet špičkových hodnot a nižší strmost kumulativní funkce poukazuje na nižší schopnost artikulace, tedy pravděpodobnou vyšší míru postižení HD.

5.2 Parametr Speech Articulation

Druhý z navržených parametrů Speech Articulation (SA) částečně vychází z parametru funkce FV, avšak zohledňuje variabilitu všech tří formantů zároveň. V rámci tohoto parametru je vypočten tzv. vektor artikulace, sklon kumulativní funkce (součtu) absolutních rozdílů hodnot po sobě jdoucích segmentů a artikulační špičkové hodnoty. Artikulační funkce je definovaná jako produkt absolutních rozdílů formantových komponentů pro interval zvolené délky (v základním nastavení: 0,5 sekund). Formantové komponenty jsou vypočítány jako mediány daných formantů podělené maximální hodnotou tohoto formantu pro daný segment. Artikulační funkce tedy popisuje jak se trojice sledovaných formantů mění v čase. Špičkové hodnoty této funkce popisují počet signifikantních změn v artikulaci pro zvolenou prahovou hodnotu, která odpovídá 1.5ti násobku průměru této artikulační funkce.

Pro výpočet tohoto parametru je stejně jako v případě parametru FV nutné získat hodnoty jednotlivých formantů pro dané segmenty a následně provést preemfázovou filtraci. Pro preemfázovou filtraci je zvolena hranice 50Hz a je zde využito Hammingova okna o délce 25ms a překryvem 50%. Signál je následně potřeba zbavit nepotřebných dat jako jsou nulové hodnoty, nekonečné hodnoty, nečíselné hodnoty a také všechny hodnoty kratší jak 3 vzorky. Následně jsou vypočteny hodnoty komponentů jakožto mediány daných formantů podělené maximální hodnotou tohoto formantu pro daný segment. Z takto vytvořených formantových komponent je poté jejich součinem poděleným celkovým počtem platných (nenulových) segmentů vytvořen vektor artikulace. Výstupem funkce je pak kromě vektoru artikulační funkce také vektor sklonu kumulativního součtu daný jako:

$$Cumsum_F = \left(\frac{\sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N |F - \mu|^2}}{\mu} \right) \cdot \left(\frac{\sum_{m=1}^N |F(n) - F(n-1)|}{\left| \vec{F} \right|} \right), \quad (5.3)$$

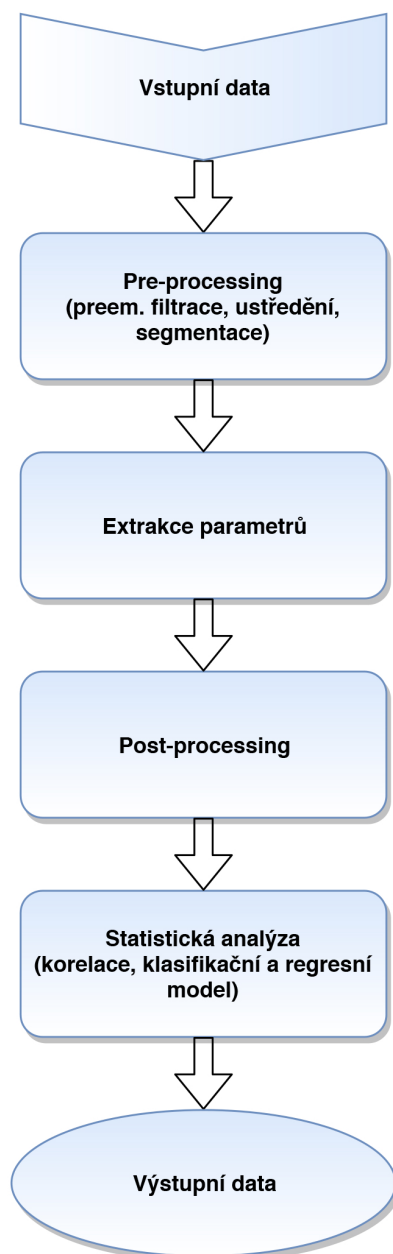
kde F je vektor hodnot daného formantu, N je počet prvků tohoto vektoru a \vec{F} je délka tohoto vektoru. μ je pak aritmetický průměr vektoru F daný vzorcem:

$$\mu = \frac{1}{N} \cdot \sum_{i=1}^N F. \quad (5.4)$$

Výpočet sklonu kumulativního součtu je tímto normalizován tak, aby se hodnoty pohybovaly pouze v rozsahu od 0 do 1. Díky tomu jsou řečníci lépe porovnatelní. V poslední fázi výpočtu jsou odstraněny veškeré prázdné hodnoty a dále je vypočítán vektor špičkových hodnot získaných stejně jako v případě parametru funkce FV viz 5.1. Stejně jako u tohoto parametru platí, že s vyššími hodnotami a strmější funkcí kumulativního součtu vykazuje řeč menší patologické příznaky.

6 REALIZACE SYSTÉMU

6.1 Blokové schéma systému



Obr. 6.1: Blokové schéma systému

6.2 Popis systému

V rámci této kapitoly je popsána funkce systému vytvořeného pro testování nově navržených parametrů. Systémem je rozuměn celkový proces zpracování a analýzy nově navržených metod, tzn. načtení testovacích dat, jejich předzpracování, výpočet parametrů, získání výstupních dat a poté jejich následná statistická analýza.

6.2.1 Načtení a zpracování dat

V první části realizace systému byla v programovém prostředí Matlab navržena funkce pro načtení zvukových nahrávek pacientů a jejich následné parametrizaci a uložení ve formátu potřebném pro další zpracování. Bylo použito databáze 53 zdravých (HP) a 99 patologických (PD) řečníků při čtení řečových cvičení standardizovaného textu (viz kap. 3.2). V rámci funkce *"demo.mat"* je provedeno načtení zvukových záznamů ve formátu .wav ze zvoleného adresáře.

6.2.2 Předzpracování

Data jsou během funkcí *"demo_articulation.mat"* a *"demo_variability.mat"* iterovány přes další navržené skripty s názvy *"formants_info.m"* a *"articulation_info.m"*, které v rámci svého cyklu využívají funkce *"praat_formants.m"*, která provádí segmentaci (viz 4.1.3) a preemfázovou filtraci (viz 4.1.2) daných signálů. Získaná výstupní jsou data jednotlivě ukládána do struktur pro každé pozorování.

6.2.3 Výpočet akustických parametrů

V rámci skriptu *"formants_info.m"* dochází k uložení formantových frekvencí jednotlivých segmentů signál pomocí funkce *"praat_formants.m"* využívající programu PRAAT. Hodnoty jsou pro obě dané funkce dále uloženy a zvlášť iterovány pomocí navržené funkce *"formant_variability.m"*. Vstupem této funkce je tak vektor již vypočtených hodnot F pro daný formant získaný pomocí volně dostupné aplikace PRAAT a také údaj o vzrokovací frekvenci daného signálu. Extrakce formantových frekvencí pro funkci *"demo_articulation.mat"* probíhá v rámci skriptu *"speech_articulation.m"*. Vstupem této funkce je předem nezpracovaný signál, vzorkovací frekvence tohoto signálu a poté údaje pro nastavení segmentace a preemfázové filtrace. V první části algoritmu jsou pomocí programu PRAAT získány hodnoty prvních tří formantů pro každý komponent funkce při využití preemfázové filtrace a segmentace dle nastavení daného ve vstupu funkce. V rámci obou funkcí jsou vytvořeny datové struktury obsahující: vektor absolutních rozdílů (*variability series*), počet špičkových hodnot (*num peaks*) a sklon kumulativní funkce (*cumsum slope*).

Výstupem pro funkci *"demo_variability.mat"* jsou tak data o změně hodnot jednotlivých formantů v čase. Z těchto dat je v rámci funkce *"formants_info.m"* vypočteno 30 dílčích akustických parametrů. Jsou to:

- sklon kumulativní funkce (*cumsum slope*)
- počet špičkových hodnot (*num peaks*)
- minimum (*min*)
- maximum (*max*)
- průměr (*mean*)
- směrodatná odchylka (STD)
- rozsah (*range*)
- 1. až 3. kvartil (Q_1 , Q_2 a Q_3)

Každý akustický parametr je takto vypočítán zvlášť pro formanty F_1, F_2, F_3 .

Výstupem pro funkci *"demo_articulation.mat"* jsou tak data o změně hodnot v čase pro první tři formanty současně. Z těchto dat je v rámci funkce *"articulation_info.m"* vypočteno 10 dílčích akustických parametrů:

- sklon kumulativní funkce (*cumsum slope*)
- počet špičkových hodnot (*num peaks*)
- minimum (*min*)
- maximum (*max*)
- průměr (*mean*)
- směrodatná odchylka (STD)
- rozsah (*range*)
- 1. až 3. kvartil (Q_1 , Q_2 a Q_3)

6.2.4 Post-processing

V závěrečné části kódu se data ve formě struktur získaná funkcí *"formants_info.m"* ukládají do tabulkového souboru *"pacienti.xlsx"*

6.3 Statistická analýza

6.3.1 Porovnání akustických parametrů

Pro porovnání rozsahu hodnot získaných akustických parametrů u HP a PD pacientů bylo použito tzv. krabicových grafů (*box plots*) pro každý ze 30 sledovaných parametrů. Funkce *"boxPlot_demo.m"* načte akustické parametry ze souboru *"pacienti_corr.xlsx"*, který kromě parametrů obsahuje také údaj o klinickém stavu pacienta pro danou nahrávku. Krabicový graf je grafickou vizualizací dat pomocí jejich kvartilů, kde šíře grafu reprezentuje míru rozptýlu sledovaných dat. Horní část grafu

je ohraničena 3. kvantilem a jeho spodní hodnota 1. Středová čára reprezentuje medián těchto hodnot. Odlehlé hodnoty jsou reprezentovány jednotlivými body v grafu. Viz příloha D.

6.3.2 Korelační analýza

Tato statistická metoda se využívá ke zjištění míry vztahu mezi jednou či více proměnnými. Dvě proměnné jsou korelované, jestliže určité hodnoty jedné proměnné mají tendenci se vyskytovat společně s určitými hodnotami druhé proměnné. Výsledná hodnota korelační analýzy, nazývaná korelační koeficient, se obvykle pohybuje v intervalu $<0,1>$, kde 0 znamená žádný vztah mezi sledovanými proměnnými, naopak 1 maximální vztah mezi nimi [22].

6.3.3 Pearsonův korelační koeficient

Pearsonův korelační koeficient r_p je pravděpodobně nejpoužívanější nástroj pro měření lineární závislosti mezi dvěma proměnnými. Pohybuje se v intervalu $<-1,1>$ a platí, že čím větší je absolutní hodnota Pearsonova korelačního koeficientu, tím vyšší je lineární závislost sledovaných proměnných. Hodnota korelačního koeficientu může být kladná a lineární se znaménkem (+), či záporná, tedy nelineární se znaménkem (-). Pro proměnné X a Y čítající n hodnot platí, že Pearsonův koeficient je roven:

$$r_s = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (6.1)$$

6.3.4 Spearmanův korelační koeficient

Tento korelační koeficient zachycuje na rozdíl od Pearsonova korelačního koeficientu (ze kterého vychází) obecně monotónní (tedy rostoucí či klesající) vztahy mezi proměnnými, nikoli pouze lineární. Jeho výhodnou vlastností je odolnost vůči zkreslení odlehlými hodnotami a také nižší výpočetní náročnost pro menší rozsahy dat. Jeho korelační koeficient r_s se pohybuje v rozsahu $<-1,1>$ a značí tzv. monotónnost dané vazby. Znaménka koeficientu zde značí jeho monotónně vzrůstající (+) či monotónně klesající vztah (-). Necht jsou hodnoty každé z proměnných X a Y seřazeny vzestupně a každé hodnotě je přiděleno pořadí. Za předpokladu, že d_i značí rozdíl ohodnocení i -tých hodnot, platí pro výpočet Spearmanova korelačního koeficientu rovnice [22]:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \quad (6.2)$$

kde n je počet hodnot proměnných. Získanou hodnotu ρ je poté potřeba porovnat s tzv. kritickou hodnotou Spearmanova korelačního koeficientu pro dané n a zvolenou chybou α . V případě, že platí $|\rho| > \rho(\alpha, n)$, je Spearmanův korelační koeficient významný na chybové hladině nejčastěji $\alpha = 0.05$, tzn. že byla prokázána vzájemná korelace zkoumaných proměnných.

6.3.5 Výpočet korelace parametrů

V rámci analýzy získaných parametrů byl vytvořen skript "correlation.m", který načte soubor "pacienti_corr.xlsx" obsahující vypočtené parametry a údaj o klinickém stavu pacienta pro dané pozorování, kde nulové hodnoty odpovídají zdravému a jednotkové hodnoty dysartrickému řečníkovi. Pomocí skriptu je pak provedena Spearmanova korelační analýza a výsledkem je tabulka ρ a p hodnot pro každý parametr. Tato analýza byla provedena za účelem zjištění nejvhodnějšího parametru pro rozlišení zdravé a dysartrické řeči. Její výsledky se nachází v tabulkách 7.1 a 7.2.

6.3.6 Strojové učení

Machine learning, neboli strojové učení (SU), patří mezi vědní disciplíny zabývající se programy a algoritmy umožňujícími strojům proces učení. Učením je rozuměno automatické zlepšování se na základě zkušeností [12]. Významným polem jeho využití je právě statistika a vyhodnocování modelů. Učící algoritmy se dají rozdělit do dvou skupin - tzv. učení s učitelem (*supervised learning*) a učení bez učitele (*unsupervised learning*). V případě učení s učitelem dodáváme algoritmu informaci o požadované výstupní hodnotě Y (učitelem je požadovaná výstupní hodnota), v opačném případě se jedná o učení bez učitele. V oblasti analýzy patologické řeči se využívá především metod učení s učitelem, zejména metody tzv. binární klasifikace a tzv. regresního modelu. V případě binární klasifikace jde zvláště o zjištění klinického stavu pacienta (zdravý/nemocný). U metody regresní analýzy je vytvořen odhad klinického stavu daného pacienta na základě rozličných škál hodnotících různé aspekty subjektu, např. progresu onemocnění apod.

V této práci bylo využito volně dostupných software pro klasifikační analýzu (CAT) a také software pro regresní analýzu (RAT).

Klasifikační analýza

V rámci klasifikační analýzy bylo využito algoritmu tzv. *random forrest*, neboli metody klasifikačních stromů (KS). Klasifikační strom je hierarchický nelineární systém

umožňující nalezení a uložení znalostí a jejich využití k analýze nových dat [12]. Při průchodu dat stromem dochází k jejich dělení v uzlech na základě podmínek. Dosažení listu klasifikačního stromu vede k predikci či klasifikaci sledovaných dat. Pomocí algoritmu byla vypočítána přesnost klasifikace na základě tzv. čtyřpolní tabulky binární klasifikace:

V rámci nastavení učicího algoritmu binární klasifikace byla použita metoda tzv. kří-

		Skutečnost	
		Pozitivní	Negativní
Predikce	Pozitivní	TP	FP
	Negativní	FN	TN

Obr. 6.2: Tabulka binární klasifikace

žové validace *cross validation*. Jedná se o metodu pro vyhodnocení přesnosti modelů strojového učení vhodnou zejména na omezeném množství testovacích dat. Data jsou vícenásobně rozdělena na trénovací a testovací data. Pomocí této obměny dochází ke zvýšení přesnosti modelu díky postupnému využití všech testovacích dat. V rámci algoritmu využitého v této práci byla aplikována tzv. *k-fold* křížová validace, která data rozděluje na *k* datových podmnožin testovacích a trénovacích dat. Tento proces se několikrát opakuje, pokaždé s jinou podmnožinou tvořící trénovací a testovací množinu dat [12]. V případě uvedeného modelu se jedná o desetinásobnou, tedy tzv. *10-fold* křížovou validaci.

Údaje jako přesnost (ACC), senzitivita (SEN), specificita (SPE) a Matthewsův korelační koeficient (MCC) jsou vypočítány na základě takto daných stavů:

TP: správně určený výskyt nemoci

TN: správně určený kontrolní řečník

FP: mylně určený výskyt nemoci

FN: mylně určený kontrolní řečník.

Přesnost:

$$ACC = (TP + TN) / (TP + TN + FP + FN) \cdot 100 [\%]. \quad (6.3)$$

Senzitivita:

$$SEN = (TP) / (TP + FN) \cdot 100 [\%]. \quad (6.4)$$

Specificita:

$$SPE = (TN) / (TN + FP) \cdot 100 [\%]. \quad (6.5)$$

Matthewsův korelační koeficient:

$$MCC = \frac{(TP \cdot TN - FP \cdot FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} [-]. \quad (6.6)$$

Přesnost vyjadřuje procentuální úspěšnost správného zhodnocení řečníka. Senzitivita vyjadřuje míru úspěšnosti zachycení nemocného řečníka v procentech, naopak specificita řečníka zdravého. Matthewsův korelační koeficient měří kvalitu binární klasifikace a jeho hodnota se pohybuje v rozsahu od -1 do 1. Pokud se MCC rovná 1, jde o bezchybnou klasifikaci, naopak pokud se rovná 0, jedná se o přesnost klasifikace, která se rovná náhodným výsledkům. Pokud by se výsledek rovnal -1, došlo by k absolutnímu rozporu mezi klasifikací a predikcí.

Regresní analýza

Pro potřebu regresní analýzy byla využita metoda "CART" (*Classification and regression trees*). Jedná se o jednu z metod rozhodovacích stromů, kde stromy vytvořené metodou CART slouží ke klasifikaci i regresi, tudíž mohou být výstupní proměnné kvalitativní i kvantitativní. Základní verze "CART" algoritmu funguje na principu nejvhodnějšího rozdělení záznamů podle jednoho z atributů [12]. Jeho předností je možnost váhování jednotlivých klasifikačních tříd na základě kterých je v rámci průchodu dat větvemi přikládána větší nebo menší váha v rámci penalizace chyby. Skutečnost, že prvek třídy A byl klasifikován jako typ B, může být penalizována odlišně, než chybné zařazení toho samého prvku do třídy C [12]. Stejně jako u klasifikační metody byla v rámci regresního modelu využita křížová validace dat se stejným nastavením 6.3.6. Kromě matice akustických parametrů byly jako vstup pro učicí model zahrnuty dodatečné údaje o klinickém stavu jednotlivých pacientů reprezentované pěticí klinických škál. Jmenovitě se jedná o škály "UPDRSIII", "UPDRSIV", "FOG", "DX", a "ACE-R". Škály "UPDRSIII" a "UPDRSIV" spadají pod sjednocený systém "UPDRS", neboli *Parkinson's disease rating scale*. Jedná se o systém škál hodnotících průběžnou progresi PN u konkrétních pacientů.

Zároveň jde o nejpoužívanější škálu pro klinickou studii PN [19]. "UPDRSIII" reprezentuje klinicky zaznamenané hodnocení motorických schopností pacienta, "UPDRSIV" průběžně nastálé komplikace v léčebné terapii. Škála "FOG", neboli *freezing of gait* popisuje závažnost výskytu tzv. *freezingu*. Jedná se poruchu chůze a motoriky vyskytující se u pacientů s PN, způsobující náhlou ztuhlost a dočasnou nemožnost pohybu. Škála ACE-R reprezentuje tzv. kognitivní deficit, tedy přechodné nebo trvalé poškození poznávacích funkcí např. výpadky paměti, pozornosti atd. Škála "DX" reprezentuje tzv. *dysartrický index* viz kap 3.1.

Pro změření přesnosti predikce regresního modelu bylo využito konvenčně široce využívaných metrik jako: *mean absolute error* (MAE), *root mean squared error* (RMSE), *mean squared error* (MSE), *estimation error rate* (EER). Metrika MAE je základní a často využívaná metrika zohledňující absolutní rozdíl mezi předpověďmi a skutečnými hodnotami. Její nevýhodou je nepřesnost způsobená faktem, že tato metrika preferuje modely s většinově převažujícími výsledky, je proto vhodné ji využívat v kombinaci s dalšími doplňujícími metrikami. Čím vyšší je hodnota MAE, tím horší je predikce. Vypočítá se jako:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (6.7)$$

Metrika MSE je vyjádřena jako:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (6.8)$$

Metrika RMSE reprezentuje odmocninu z průměru mocniny predikční chyby, z toho plyne, že se snižujícím se RMSE se také snižuje přesnost predikce. Je definována jako:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (6.9)$$

Metrika EER se využívá pro stanovení predikční chyby. Představuje modifikovanou metriku MAE rozšířenou o vazbu k rozsahu hodnot předpovídané proměnné. Její hodnotu lze vypočítat jako:

$$EER = \frac{1}{n \cdot r} \sum_{i=1}^n |y_i - \hat{y}_i| \cdot 100[\%]. \quad (6.10)$$

Proměnná n reprezentuje počet pravdivých/předpovězených hodnot, y_i a \hat{y}_i představují skutečné a modelem předpovězené proměnné a hodnota r značí rozsah hodnot daný konkrétní klinickou škálou využitou pro danou předpověď.

Výsledky testování navržených algoritmů pomocí strojového učení a také korelační analýza jejich akustických parametrů jsou uvedeny v kapitole 7.

7 VÝSLEDKY

7.1 Korelační analýza

Korelační analýza byla provedena za účelem zjištění míry korelace dílčích akustických parametrů obou navržených algoritmů s klinickým stavem daných pacientů. Jelikož se v případě klinického stavu pacientů jedná o kategorickou veličinu, pro výpočet korelace byl využit Spearmanův korelační koeficient (viz kap. 6.3.4) s hladinou významnosti $\alpha = 0.05$. Jak lze vidět v tabulce 7.1, pro parametr FV bylo dosaženo nejvyšší míry korelačního koeficientu ρ pro akustický parametr "F1_slope" (tedy sklon kumulativní funkce) při hodnotě p mírně přesahující hladinu významnosti. Nejvyšších záporných hodnot korelačního koeficientu, tedy nejvyšších nepřímo úměrných hodnot mezi korelovanými veličinami bylo dosaženo u akustických parametrů "F1_num" a "F2_num" (tedy počtu špičkových hodnot prvního a druhého formantu) s hodnotami p nepřesahujícími stanovenou hladinu významnosti. Záporný korelační koeficient u těchto akustických parametrů poukazuje na snížení počtu výrazných špičkových hodnot prvních dvou formantů u pacientů s HD.

V tabulce 7.2 se nachází výsledky korelační analýzy pro parametr SA. Zde můžeme vidět nejvýznamnější výsledky pro akustické parametry "F_num" a "F_min", kde hodnoty v rámci hladiny významnosti dosahují nejvyšších záporných, tedy nepřímo úměrných hodnot korelačního koeficientu ρ . Stejně jako u parametru FV tento výsledek poukazuje na snížení počtu špičkových hodnot a celkového minima hodnot prvních tří formantů u pacientů s HD, z čehož vyplývá snížení hybnosti artikulačních orgánů způsobených HD.

Výsledky korelační analýzy

	F1_slope [-]	F2_slope [-]	F3_slope [-]	F1_num [-]	F2_num [-]	F3_num [-]	F1_min [-]	F2_min [-]	F3_min [-]	F1_max [-]
ρ [-]	0,1505	0,0589	0,0885	-0,1982	-0,2033	-0,1183	0,0036	-0,0640	0,0253	-0,0838
p [-]	0,0641	0,4703	0,2778	0,0143	0,0120	0,1463	0,9647	0,4287	0,7567	0,3044
	F2_max [-]	F3_max [-]	F1_mean [-]	F2_mean	F3_mean [-]	F1_std [-]	F2_std [-]	F3_std [-]	F1_range [-]	F2_range [-]
ρ [-]	0,1288	-0,0281	-0,0029	0,0045	0,0549	-0,0662	0,0426	0,0961	-0,0851	0,1313
p [-]	0,1136	0,7305	0,9708	0,9555	0,5017	0,4175	0,6020	0,2388	0,2971	0,1067
	F3_range [-]	F1_Q1 [-]	F1_Q2 [-]	F1_Q3 [-]	F2_Q1 [-]	F2_Q2 [-]	F2_Q3 [-]	F3_Q1 [-]	F3_Q2 [-]	F3_Q3 [-]
ρ [-]	-0,0281	0,0070	0,0158	0,0558	-0,0898	-0,0718	-0,0360	-0,0089	0,0180	0,0410
p [-]	0,7305	0,9310	0,8459	0,4943	0,2710	0,3787	0,6594	0,9126	0,8249	0,6154

Tab. 7.1: Korelační analýza parametru Formant Variability.

	F_slope [-]	F_num [-]	F_min [-]	F_max [-]	F_mean [-]	F_std [-]	F_range [-]	F_Q1 [-]	F_Q2 [-]	F_Q3 [-]
ρ [-]	0,0117	-0,2141	-0,1732	-0,0162	-0,0368	-0,0080	-0,0162	-0,0193	-0,0303	-0,1013
p [-]	0,8852	0,0080	0,0328	0,8429	0,6525	0,9218	0,8429	0,8129	0,7103	0,2142

Tab. 7.2: Korelační analýza parametru Speech Articulation.

7.2 Klasifikační analýza

Klasifikační analýza byla realizována zvlášť pro oba navrhované algoritmy a jejich dílčí akustické parametry. Pro každý z těchto algoritmů byla provedena tzv. *univariable analysis*, tedy analýza přesnosti metody pomocí jedné proměnné, respektive jednoho testovaného akustického parametru. Dále byla provedena tzv. *multivariable analysis*, tedy analýza s pomocí více proměnných, v tomto případě s pomocí všech získaných akustických parametrů pro daný algoritmus. V rámci metody strojového učení bylo využito algoritmu *random forrest* při tzv. *k-fold cross validace* s deseti opakováními viz 6.3.6

Výsledky pro klasifikační analýzu v tomto textu obsahují vybrané akustické parametry s nejvyšší hodnotou MCC (viz kap. 6.3.6).

Klasifikační analýza s jednou proměnnou

Pro parametr FV bylo dosaženo nejpřesnější predikce pro akustický parametr "F1_min" (SEN mean=0,66/std=0,15), tedy minima prvního formantu viz tab. 7.3.

	MCC mean	MCC std	ACC mean	ACC std	SEN mean	SEN std	SPE mean	SPE std
F1_min	0,2741	0,2452	0,6502	0,1185	0,6681	0,1546	0,6131	0,2360
F3_Q3	0,2143	0,2494	0,6246	0,1178	0,6666	0,1640	0,5529	0,2310
F2_num	0,1419	0,2404	0,5881	0,1182	0,6337	0,1629	0,5152	0,2216
F3_Q2	0,1204	0,2527	0,5852	0,1125	0,6447	0,1493	0,4815	0,2536
F3_std	0,1180	0,2290	0,5804	0,1086	0,6142	0,1601	0,5101	0,2301

Tab. 7.3: Výsledky klasifikační analýzy s jednou proměnnou pro parametr FV.

U parametru funkce SA bylo dosaženo nejpřesnějších predikcí u akustického parametru "F_slope" (SEN mean=0,60/std=0,16), tedy sklonu kumulativní funkce prvních tří formantů viz tab. 7.4.

	MCC mean	MCC std	ACC mean	ACC std	SEN mean	SEN std	SPE mean	SPE std
F_slope	0,1088	0,2540	0,5631	0,1187	0,6020	0,1602	0,5151	0,2603
F_num	0,0848	0,2160	0,5689	0,1109	0,6339	0,1885	0,4488	0,2365
F_max	0,0722	0,2678	0,5543	0,1298	0,6082	0,1671	0,4696	0,2351
F_range	0,0647	0,2638	0,5519	0,1327	0,5919	0,1779	0,4699	0,2323
F_mean	0,0502	0,2495	0,5394	0,1208	0,5693	0,1612	0,4817	0,2237

Tab. 7.4: Výsledky klasifikační analýzy s jednou proměnnou pro parametr SA.

Klasifikační analýza s více proměnnými

Pro potřebu komplexnějšího zhodnocení akustických parametrů u parametrů FV i SA byla provedena také klasifikační analýza s více, respektive všemi dostupnými proměnnými (akustickými parametry) a její výsledky jsou uvedeny v tabulkách 7.5 a 7.6.

	MCC mean	MCC std	ACC mean	ACC std	SEN mean	SEN std	SPE mean
Parametry	-0,0710	0,2528	0,4854	0,1292	0,5380	0,1743	0,3908

Tab. 7.5: Výsledky klasifikační analýzy se všemi parametry pro parametr FV.

	MCC mean	MCC std	ACC mean	ACC std	SEN mean	SEN std	SPE mean
Parametry	0,1141	0,2586	0,5725	0,1226	0,6103	0,1706	0,5051

Tab. 7.6: Výsledky klasifikační analýzy se všemi parametry pro parametr funkce SA.

U parametru FV (viz tab. 7.5) bylo v případě klasifikační analýzy s více parametry dosaženo mírně záporné hodnoty MCC, tedy hodnoty přiklánějící se k celkové chybné predikci. V případě parametru algoritmu SA (viz tab. 7.6) bylo získáno lepších výsledků, tedy celkové průměrné hodnoty MCC vyšší než 0,1.

7.3 Regresní analýza

Regresní analýza byla stejně jako klasifikační realizována zvlášť pro oba navrhované parametry a jejich dílčí akustické parametry. Pro každý z těchto parametrů byla taktéž provedena analýza přesnosti metody s jednou i více proměnnými (*univariable* a *multivariable* analýza). V rámci metody strojového učení bylo využito algoritmu *CART* při tzv. *k-fold cross validaci* s deseti opakováními viz 6.3.6. Vzhledem k povaze testování regresního modelu bylo třeba získat výsledky na základě aplikace každé z uvedených klinických škál viz kap. 6.3.6. V případě regresní analýzy bylo tedy využito pouze pozorování pacientů s HD pro zhodnocení přesnosti modelu na základě hodnot daných klinických škál pro jednotlivé pacienty.

Regresní analýza s jednou proměnnou - UPDRSIII

V případě regresní analýzy byla jako primární metrika hodnocena hodnota střední absolutní chyby MAE (viz kap. 6.3.6). Nejnižší hodnoty této chyby pro regresní analýzu s jednou proměnnou pro funkci FV a "UPDRSIII" dosáhl akustický parametr

"F2_range" (MAE mean=9,96/std=2,55), tedy rozsah hodnot druhého formantu viz tab. 7.7.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F2_range	9,9683	2,5574	19,1697	4,9180	160,8408	77,9780	12,3626	2,9828
F2_Q3	10,4867	1,1495	20,1667	2,2105	163,6274	51,7154	12,6668	1,8796
F2_max	10,7150	2,9361	20,6057	5,6463	178,9131	80,3385	13,0157	3,2495
F2_min	10,9801	3,5386	21,1156	6,8050	206,4692	99,2168	13,9223	3,7473
F3_range	11,1267	3,4672	21,3974	6,6676	193,2135	113,5791	13,4083	3,8629

Tab. 7.7: Výsledky regresní analýzy s jednou proměnnou dle škály UPDRSIII pro parametr FV.

Naproti tomu u parametru SA šlo o akustický parametr "F_num" (MAE mean=10,2/std=2,06), tedy celkový počet špičkových hodnot prvních tří formantů viz tab. 7.8.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F_num	10,2475	2,0612	19,7068	3,9639	151,0329	51,7555	12,1059	2,2309
F_slope	11,0388	2,7919	21,2284	5,3691	192,3768	89,6437	13,4410	3,6079
F_range	11,6226	2,3203	22,3511	4,4622	212,1073	80,9382	14,3090	2,8595
F_mean	11,8462	3,9946	22,7812	7,6820	229,2983	183,9470	14,2712	5,3367
F_std	12,1477	3,3343	23,3609	6,4121	231,0567	99,4380	14,8927	3,2083

Tab. 7.8: Výsledky regresní s jednou proměnnou dle škály UPDRSIII pro parametr SA.

Regresní analýza s jednou proměnnou - UPDRSIV

V případě regresní analýzy funkce FV pro škálu "UPDRSIV" bylo dosaženo nejnižších hodnot MAE pro akustický parametr "F2_range" (MAE mean=2,44/std=0,77) podobně jako v případě predikce u škály "UPDRSIII" viz tab. 7.9.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F2_range	2,4422	0,7716	27,1360	8,5728	9,6789	4,9692	2,9991	0,8719
F3_num	2,4607	0,4808	27,3406	5,3425	9,2628	3,7860	2,9850	0,6256
F3_slope	2,5262	0,3407	28,0684	3,7860	8,4939	2,4873	2,8846	0,4388
F2_slope	2,5445	0,4797	28,2721	5,3305	9,3193	3,2525	3,0055	0,5638
F1_std	2,5624	0,6226	28,4713	6,9174	9,8441	4,2519	3,0640	0,7116

Tab. 7.9: Výsledky regresní analýzy s jednou proměnnou dle škály UPDRSIV pro parametr FV.

Podobná predikce vyšla také pro analýzu parametru SA s nejpřesnějším výsledkem predikce pro "F_num" (MAE mean=2,36/std=0,47) viz tab. 7.10.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F_num	2,3656	0,4703	26,2844	5,2254	7,9234	2,5900	2,7769	0,4858
F_min	2,5589	0,4584	28,4326	5,0933	10,2661	4,2713	3,1415	0,6644
F_Q2	2,6084	0,5921	28,9826	6,5792	10,1643	3,7241	3,1209	0,6864
F_mean	2,6826	0,6280	29,8062	6,9774	12,0337	4,9228	3,4066	0,6900
F_slope	2,7006	0,6539	30,0072	7,2653	10,6047	4,9175	3,1840	0,7204

Tab. 7.10: Výsledky regresní analýzy s jednou proměnnou dle škály UPDRSIV pro parametr SA.

Regresní analýza s jednou proměnnou - FOG

V případě regresní analýzy s jednou proměnnou a škálou "FOG" pro parametr FV bylo predikováno nejpřesnějších výsledků u akustického parametru "F1_Q2" (MAE mean=4,76/std=1,14), tedy hodnoty mediánu hodnot prvního formantu viz tab. 7.11.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F1_Q2	4,7619	1,1437	23,8096	5,7185	36,4642	18,7414	5,8784	1,4562
F2_std	5,0846	1,2697	25,4229	6,3486	40,6168	19,3004	6,2253	1,4386
F3_num	5,2638	0,9520	26,3189	4,7600	39,7198	13,9960	6,2233	1,0490
F2_mean	5,4680	1,1056	27,3400	5,5280	44,7602	14,6978	6,6065	1,1129
F1_Q1	5,5679	1,3547	27,8396	6,7734	45,1333	19,2266	6,5507	1,5710

Tab. 7.11: Výsledky regresní analýzy s jednou proměnnou dle škály FOG pro parametr FV.

Pro parametr SA v případě škály "FOG" byl svými hodnotami nejpřesnější akustický parametr "F_num" (MAE mean=4,73/std=0,72) podobně jako v případě škál "UPDRSIII" a "UPDRSIV" viz tab. 7.12.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F_num	4,7378	0,7289	23,6889	3,6443	31,0908	11,3198	5,5011	0,9594
F_max	5,3059	1,0229	26,5293	5,1145	45,4683	17,0948	6,6331	1,2782
F_min	5,5088	1,4503	27,5440	7,2515	45,1836	22,6284	6,5302	1,6801
F_Q1	5,5554	1,3584	27,7770	6,7922	45,5670	20,2131	6,6064	1,4614
F_std	5,5812	0,9648	27,9058	4,8241	46,4639	17,7386	6,7078	1,2779

Tab. 7.12: Výsledky regresní analýzy s jednou proměnnou dle škály FOG pro parametr SA.

Regresní analýza s jednou proměnnou - DX

Pro regresní analýzu s jednou proměnnou u parametru FV a škálu DX s nejnižší hodnotou MAE vyšel akustický parametr "F1_max" (MAE mean=6,56/std=2,42), tedy maximální hodnota prvního parametru viz tab. 7.13.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F1_max	6,5658	2,4261	11,5190	4,2562	84,2234	76,0474	8,4389	3,8018
F3_num	6,9042	1,7022	12,1127	2,9862	84,6565	49,5639	8,8274	2,7351
F1_std	6,9555	2,7066	12,2026	4,7484	110,7745	97,5781	9,6068	4,5319
F1_range	7,1146	2,2788	12,4818	3,9979	90,3784	69,0855	8,9683	3,3246
F3_Q2	7,2872	2,7246	12,7846	4,7800	98,2978	96,9001	9,1577	4,0048

Tab. 7.13: Výsledky regresní analýzy s jednou proměnnou dle škály DX pro parametr FV.

V rámci parametru funkce SA a škály "DX" vyšel s nejnižší chybou akustický parametr "F_range" (MAE mean=6,74/std=2,12), tedy rozsah hodnot prvních tří parametrů viz tab. 7.14.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F_range	6,7407	2,1292	11,8258	3,7355	91,8901	81,4480	8,9533	3,6100
F_num	6,7921	2,6641	11,9160	4,6738	86,3393	94,3471	8,4723	4,0220
F_Q1	6,8991	1,6555	12,1037	2,9043	91,9508	58,0218	9,2305	2,7383
F_slope	7,1577	2,7543	12,5574	4,8321	97,2314	93,5395	9,0725	4,0716
F_max	7,2227	3,5485	12,6714	6,2254	109,8644	118,1671	9,4134	4,8593

Tab. 7.14: Výsledky regresní analýzy s jednou proměnnou dle škály DX pro parametr SA.

Regresní analýza s jednou proměnnou - ACE-R

Regresní analýza s jednou proměnnou pro parametr funkce FV a škálu "ACE-R" ukázala "F2_range" (MAE mean=6,41/std=1,24) jakožto nejpřesnější z akustických parametrů podobně jako u škál "UPDRSIII" a "UPDRSIV" viz tab. 7.15.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F2_range	6,4193	1,2483	16,0483	3,1208	69,6970	25,9157	8,1971	1,6684
F1_num	7,2202	1,7594	18,0505	4,3984	86,3426	41,5222	8,9981	2,4442
F2_Q3	7,2729	2,0118	18,1822	5,0295	91,5424	51,7201	9,2390	2,6211
F2_max	7,3958	1,7269	18,4895	4,3172	97,8479	49,4609	9,6015	2,5074
F1_Q3	7,4820	2,4314	18,7050	6,0785	98,9279	74,6804	9,3884	3,4617

Tab. 7.15: Výsledky regresní analýzy s jednou proměnnou dle škály ACE-R pro parametr FV.

Pro škálu "ACE-R" a parametr SA se ukázal jako nejpřesnější akustický parametr "F_num" (MAE mean=6,76/std=1,83), tedy počet špičkových hodnot prvních tří formantů. Podobně tomu bylo u škály "FOG", "UPDRSIII" a "UPDRSIV" viz tab. 7.16.

	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
F_num	6,7682	1,8327	16,9204	4,5817	72,2534	47,3564	8,1544	2,5297
F_max	7,4628	1,7694	18,6570	4,4234	84,5881	32,4530	9,0496	1,7300
F_Q1	7,7140	1,4674	19,2849	3,6685	88,7943	33,3560	9,2617	1,8303
F_range	7,7158	1,6629	19,2895	4,1573	91,5936	39,5280	9,3607	2,1004
F_Q2	7,8100	1,9733	19,5250	4,9333	100,1996	45,6130	9,7510	2,3844

Tab. 7.16: Výsledky regresní analýzy s jednou proměnnou dle škály ACE-R pro parametr SA.

Regresní analýza s více proměnnými

V obou případech regresní analýzy parametrů FV a SA s použitím všech dostupných akustických parametrů bylo dosaženo nejpřesnějších výsledků, tedy nejnížší hodnoty MAE pro škálu "UPDRSIV" s hodnotami $MAE \approx 2,5$. Relativně přesných predikcí bylo dosaženo také u škály "FOG" s hodnotami $MAE \approx 5,5$. Naopak predikcí s největší chybou bylo v obou případech dosaženo s použitím škály "UPDRSIII" s $MAE > 12$. Výsledky obou analýz se nachází v tabulkách 7.17 a 7.18.

Škála	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
DXcelkem	7,9058	3,3118	13,8698	5,8103	127,0264	123,9082	10,2630	4,9099
ACE-R	8,6823	2,0928	21,7058	5,2320	118,0124	69,7585	10,4642	3,0754
UPDRSIII	12,067	3,8364	23,2068	7,3778	229,3870	130,7261	14,5745	4,3420
FOG	5,8580	1,5172	29,2902	7,5864	51,2873	23,2241	7,0119	1,5347
UPDRSIV	2,6475	0,4779	29,4169	5,3105	10,5482	3,1033	3,2156	0,4805

Tab. 7.17: Výsledky regresní analýzy se všemi parametry dle jednotlivých škál pro parametr FV.

Škála	MAE mean	MAE std	EER mean	EER std	MSE mean	MSE std	RMSE mean	RMSE std
DXcelkem	7,1743	3,1664	12,5865	5,5552	109,7546	131,4447	9,4222	4,8276
ACE-R	7,6709	2,3455	19,1773	5,8639	96,3858	46,1252	9,5788	2,2683
UPDRSIII	12,2527	2,7797	23,5630	5,3456	236,8504	113,5327	14,9624	3,7972
FOG	5,1109	1,7735	25,5545	8,8678	43,3408	25,4023	6,3417	1,8628
UPDRSIV	2,4749	0,8283	27,4999	9,2036	10,6118	7,3086	3,0959	1,0680

Tab. 7.18: Výsledky regresní analýzy se všemi parametry dle jednotlivých škál pro parametr funkce SA.

8 ZÁVĚR

Tato bakalářská práce se zabývá neinvazivní, objektivní analýzou Parkinsonovy nemoci na základě akustické analýzy hypokinetické dyzartrie pomocí moderních metod číslicového zpracování řečových signálů pro účely tzv. parametrizace řeči. Parametrizace řeči kvantifikuje jednotlivé patologické aspekty hypokinetické dyzartrie v obas-tech: fonace, artikulace a prozodie řeči. V rámci bakalářské práce byla nastudována a podrobně popsána problematika Parkinsonovy nemoci a hypokinetické dyzartrie (následující subsystémy tvorby řeči: fonace, artikulace, prozodie řeči). Konkrétně byla věnována pozornost zejména poškození artikulace. V práci byly následně pro-studovány, popsány a také implementovány konvenční metody předzpracování dy-zartrické řeči: odstranění stejnosměrné složky, preemfázová filtrace a tzv. segmentace řeči. Dále byla nastudována, popsána a také implementována tzv. konvenční para-metrizace dyzartrické artikulace: lineární predikční analýza, formantové kmitočty a z nich odvozené parametry, které jsou běžně používány pro popis pohyblivosti mluvidel (VSA, FCR, VAI).

Kromě nastudování konvenčních metod parametrizace byly také navrženy, im-plementovány a otestovány zcela nové parametry pro kvantifikaci řeči hodnotící dy-sartrickou artikulaci. Všechny metody byly implementovány v programovém pro-středí Matlab. Pro účely kvantifikace řeči byl využit řečový korpus PARCZ, který byl vytvořen ve spolupráci výzkumníků z Fakulty Elektrotechniky a Komunikač-ních Technologií Vysokého Učení Technického v Brně a Lékařské fakulty Masary-kovy univerzity a Fakultní nemocnice u sv. Anny v Brně. Přesnost nově navrže-ných parametrů byla potom hodnocena pomocí tzv. korelační analýzy a také metod strojvého učení, které byly v rámci této práce nastudovány a popsány. Za účelem statistické analýzy výsledků nově implementovaných parametrů pomocí strojvého učení bylo dále využito tzv. klasifikační a regresní analýzy. V rámci regresní analýzy bylo využito konvenčně užívaných klinických škál jako "UPDRSIII", "UPDRSIV", "FOG", "ACE-R" a "DX" za účelem predikce klinického stavu řečníka daného tě-mito škálami. Dílčí akustické parametry byly také porovnány pro jak zdravou, tak dysartrickou řeč pomocí krabicových diagramů.

V případě korelační analýzy se u parametru Formant Variability (FV) uká-zal jako nejpřesnější akustický parametr "F1_slope", tedy parametr popisující str-most funkce kumulativního součtu významných změn prvního formantu. Nižší str-most kumulativního součtu koresponduje se sníženou pohyblivostí artikulačních or-gánů typickou pro pacienty s hypokinetickou dysartrií. Záporných hodnot korelace, tedy nepřímé úměry bylo dosaženo zvláště u akustických parametrů "F1_num" a "F2_num", které poukazují na nižší počet špičkových hodnot prvního a druhého formantu ve sledované řeči, tedy projev snížení schopnosti artikulace daných řeč-

níků. Stejný jev byl pozorován také pro výsledky korelační analýzy funkce SA, která zohledňuje chování všech tří formantů současně. Pro parametr FV popisující variabilitu jednotlivých formantů v čase bylo při klasifikační analýze s jedním parametrem dosaženo nejpřesnější predikce pro dílčí akustický parametr "F1_min" při senzitivitě 66%. To poukazuje na celkové snížení minimálních změn prvního formantu, tedy na zhoršení schopnosti artikulace. Pro funkci Speech Articulation (SA) se ukázal jako nejpřesnější akustický parametr "F_slope" při senzitivitě 60%. Tento parametr popisující celkovou změnu všech tří formantů v čase tak nabíral nejpřesnějších predikcí pro akustický parametr popisující strmost kumulativního součtu hodnot rozdílů po sobě jdoucích formantů. Čím vyšší je v tomto případě strmost funkce kumulativního součtu, tím nižší je predikovaný vliv HD.

Při regresní analýze parametru FV dle uvedených škál bylo dosaženo nejlepších výsledků především pro akustický parametr "F2_range", který reprezentuje variabilitu rozdílu mezi maximální a minimální hodnotou druhého formantu. Snížení celkového rozdílu minim a maxim těchto formantů značí celkové snížení artikulačních schopností daných menším množstvím výrazných formantových výkyvů ve sledované řeči. Pro parametr SA se jako nejpřesnější ukázal akustický parametr "F_num", který stejně jako v případě korelační analýzy poukazuje na nižší počet špičkových hodnot v signálu, tedy projev snížení schopnosti artikulace daných řečníků. V obou případech regresní analýzy parametrů FV a SA s použitím všech dostupných akustických parametrů bylo dosaženo nejpřesnějších výsledků, tedy nejnižší hodnoty MAE pro škálu "UPDRSIV" s hodnotami $MAE \approx 2,5$.

V budoucnu by na tuto práci mohlo být navázáno navržením dalších parametrů se zaměřením na ostatní oblasti řeči (prozodie, intonace atd.) či vytvořením parametrů kombinujících již navržené pro potřebu robustnější analýzy. Tyto parametry by dále bylo možné porovnat s konvenčně využívanými a také testovat pro odhad dalších klinických škál.

LITERATURA

- [1] BOERSMA, P.; WEENIK, D. *Praat: doing phonetics by computer [Computer program]* [online]. Version 5.3.51, retrieved 2 June 2013 [cit. 22. 5. 2015]. Dostupné z URL: <<http://www.fon.hum.uva.nl/praat/>>
- [2] CAEKEBEKE, J. F.; JENNEKENS-SCHINKEL, A.; VAN DER LINDEN, M.E.; BURUMA, O. J.; ROOS, R. A. *The Interpretation of Dysprosody in Patients with Parkinson's Disease* [online]. J Neurol Neurosurg Psychiatry, ročník 54, č.2, 1991: s. 145–148 [cit. 23. 11. 2014]. Dostupné z URL:<<http://jnnp.bmj.com/content/54/2/145.full.pdf>>
- [3] CANTER G. J. *Speech Characteristics of Patients with Parkinson's Disease: III. Articulation, Diadochokinesis, and Over-All Speech Adequacy* [online]. J Speech Hear Disord, ročník 30, č. 3, 1965: s. 217–224 [cit. 22. 11. 2014]. Dostupné z URL: <<http://jshd.asha.org/cgi/reprint/30/3/217.pdf>>
- [4] DEJONCKERE, P. H. *Assessment of Voice and Respiratory Function* [online]. Surgery of Larynx and Trachea, editace M. Remacle; H. E. Eckel, Springer Berlin Heidelberg, 2010, ISBN 978-3-540-79135-5, s. 11–26. [cit. 17. 11. 2014]. Dostupné z URL: <http://dx.doi.org/10.1007/978-3-540-79136-2_2>
- [5] FLETCHER, S. G. *Time-by-count Measurement of Diadochokinetic Syllable Rate* [online]. J Speech Hear Res, ročník 15, č. 4, 1972: s.763–770 [cit. 23. 11. 2014]. Dostupné z URL:<<http://jslhr.pubs.asha.org/article.aspx?articleid=1768168>>
- [6] GOBERMAN, A. M. *Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson Disease* [online]. Med Sci Monit, ročník 11, č. 3, 2005: s. CR109–116 [cit. 23. 11. 2014]. Dostupné z URL: <<http://www.ncbi.nlm.nih.gov/pubmed/15735562>>
- [7] GOBERMAN, A. M.; COELHO, C. *Acoustic analysis of parkinsonian speech I: speech characteristics and L-Dopa therapy* [online]. NeuroRehabilitation, ročník 17, č. 3, 2002: s. 237–246 [cit. 22. 11. 2014]. Dostupné z URL: <<http://www.ncbi.nlm.nih.gov/pubmed/12237505>>
- [8] HARTELIUS L.; SVENSSON P. *Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: a survey.* Folia Phoniatr Logop, ročník 46, č. 1, 1994: s. 9–17
- [9] HAZAN, H.; HILU, D.; MANEVITZ, L.; RAMIG, L.; SAPIR, S. *Early diagnosis of Parkinson's disease via machine learning on speech data* In Electrical

- Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of, 2012, s. 1–4.
- [10] HILLENBRAND, J.; GETTY, L. A.; CLARK, M. J.; WHEELER, K. *Acoustic Characteristics of American English Vowels* [online]. J Acoust Soc Am, ročník 97, č. 5, 1995: s. 3099–3111. [cit. 17. 11. 2014]. Dostupné z URL: <<http://talkingneanderthal.com/HillenbrandGettyClarkWheeler.pdf>>.
 - [11] HILKER R.; SCHWEITZER K.; COBURGER S.; GHAEMI M.; WEISENBACH S.; JACOBS A. *Nonlinear progression of Parkinson disease as determined by serial positron emission tomographic imaging of striatal fluorodopa F 18 activity* Arch Neurol, ročník 62, č. 3, 2005: s.378–382.
 - [12] HONZÍK P. *Strojové učení*. Brno: Vysoké učení technické v Brně, 2006
 - [13] KENT, R. D.; KENT, J. F.; DUFFY, J.; WEISMER, G. *The Dysarthrias: Speech-voice Profiles, Related Dysfunctions, and Neuropathology*. Journal of Medical Speech-Language Pathology, ročník 6, č. 4, 1998: s. 165–211.
 - [14] MEKYSKA, J.; REKTOROVA, I.; SMEKAL, Z. *Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease*. [online]. Telecommunications and Signal Processing (TSP), 2011 34th International Conference on, 2011: s. 408–412. [cit. 25. 11. 2016]. Dostupné z URL: <http://dx.doi.org/10.1109/TSP.2011.6043700>.
 - [15] MEKYSKA, J.; SMÉKAL, Z.; KOŠTÁLOVÁ, M.; MRAČKOVÁ, M.; SKUTILOVÁ, S.; REKTOROVÁ, I. *Motor Aspects of Speech Impairment in Parkinson's Disease and their Assessment*. Cesk Slov Neurol N, ročník 74, č. 6, 2011: s. 662–668.
 - [16] METTER, E. J.; HANSON, W. R. *Clinical and Acoustical Variability in Hypokinetic Dysarthria* [online]. J Commun Disord, ročník 19, č. 5, 1986: s. 347–366 [cit. 22. 11. 2014]. Dostupné z URL: <<http://www.ncbi.nlm.nih.gov/pubmed/3490498>>
 - [17] PALKOVÁ, Z. *Fonetika a fonologie češtiny s obecným úvodem do problematiky oboru* Karolinum, Praha 1994. ISBN 80-7066-843-1
 - [18] PSUTKA, J., et al. *Mluvíme s počítačem česky* Praha: ACADEMIA, 2006.752 s. ISBN 80-200-1309-1.
 - [19] RAMAKER, C.; MARINUS, J.; STIGGELBOUT, A.; HILTER, B. *Systematic evaluation of rating scales for impairment and disability in Parkinson's*

- disease: Movement Disorders..* 2002: s. 867–876. Dostupné z URL: <<https://onlinelibrary.wiley.com/doi/full/10.1002/mds.10248>>
- [20] SAPIR, S.; RAMIG, L. O.; SPIELMAN, J. L.; FOX, C. *Formant Centralization Ratio (FCR): A proposal for a new acoustic measure of dysarthric speech* [online]. J Speech Lang Hear Res, ročník 53, č. 1, 2010: s. 1–20. [cit. 17. 11. 2014]. Dostupné z URL: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2821466/>>
 - [21] SAPIR, S.; RAMIG, L.; FOX, C. *Speech and Swallowing Disorders in Parkinson Disease* [online]. Curr Opin Otolaryngol Head Neck Surg, ročník 16, č. 3, 2008: s. 205–210 [cit. 17. 11. 2014]. Dostupné z URL: <<http://www.ncbi.nlm.nih.gov/pubmed/8956101>>
 - [22] SHESKIN, D. *Handbook of Parametric and Nonparametric Statistical Procedures: Second Edition*. Chapman and Hall/CRC, 2000. 2nd Edition. ISBN 978-1-58488-133-9.
 - [23] SKODDA, S.; VISSER, W.; SCHLEGEL, U. *Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease*. J Neural Transm, ročník 117, 2010: s. 197–205.
 - [24] SMĚKAL, Z. *Číslicové zpracování řeči*. Brno: Vysoké učení technické v Brně, 2009: s. 122.
 - [25] ZWIRNER, P.; BARNES, G. J. *Vocal Tract Steadiness: a Measure of Phonatory and Upper Airway Motor Control during Phonation in Dysarthria* [online]. J Speech Hear Res, ročník 35, č. 4, 1992: s. 761–768 [cit. 22. 11. 2014]. Dostupné z URL:<<http://jslhr.pubs.asha.org/article.aspx?articleid=1779189>>

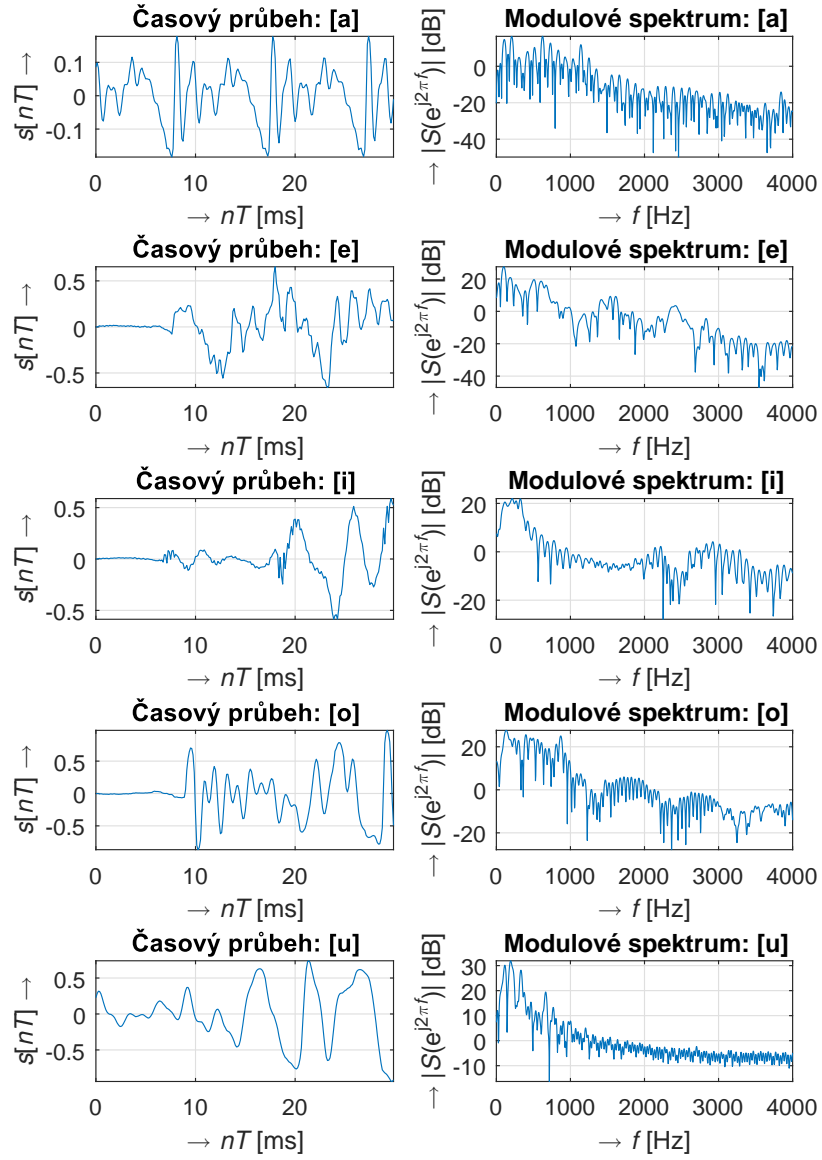
SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

α	hladina významnosti korelace
ACC	accuracy - přesnost
DDK	diadochokinetic tasks - diadochokinetické úlohy
EER	estimated error rate - odhadovaná míra chyb
F_1	frekvence prvního formantu
F_2	frekvence druhého formantu
F_3	frekvence třetího formantu
F_s	vzorkovací kmitočet
FV	parametr Formant Variability
HD	hypokinetická dysartrie
IQR	inter quartile range - rozdíl 3. a 1. kvartilu
KS	klasifikační strom
LPC	linear predictive coding - lineární predikční analýza
MAE	mean absolute error - střední absolutní chyba
MCC	Matthewsův korelační koeficient
MSE	mean squared error - střední kvadratická chyba
r_s	Parsonův korelační koeficient
p	p-hodnota korelace
PD/PN	Parkinson's disease - Parkinsonova nemoc
ρ	Spearmanův korelační koeficient
SA	parametr Speech Articulation
SEN	sensitivity - senzitivita
SPE	specificity - specifická
SU	strojové učení
VAI	Vowel Articulation Index
VSA	Vowel Space Area
Q_1	hodnota 1. kvartilu
Q_2	hodnota 2. kvartilu (medián)
Q_3	hodnota 3. kvartilu

SEZNAM PŘÍLOH

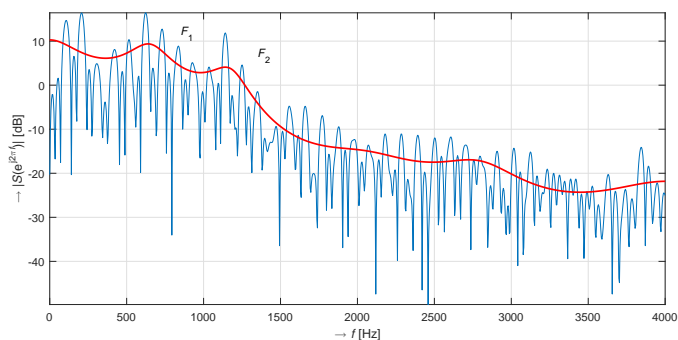
A	Fonace českých samohlásek	51
B	Formanty českých samohlásek	52
C	Zdrojové kódy parametrů	54
C.1	Kód parametru funkce Formant Variability	54
C.2	Kód parametru funkce Speech Articulation	55
D	Porovnání parametrů	58
D.1	Formant variability	58
D.2	Speech articulation	63
E	Obsah přiloženého CD	65

A FONACE ČESKÝCH SAMOHLÁSEK

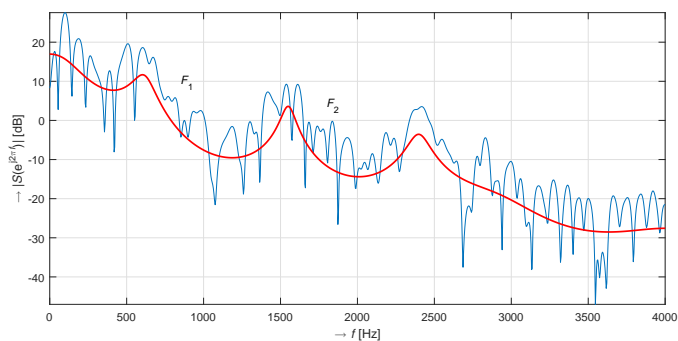


Obr. A.1: Časový průběh a jednostranné modulové spektrum fonace pěti českých samohlásek: [a], [e], [i], [o], [u]. Parametry: $Fs = 8000$ Hz, $NFFT = 2048$. Signál každé samohlásky je omezen pouze na prvních 30 ms.

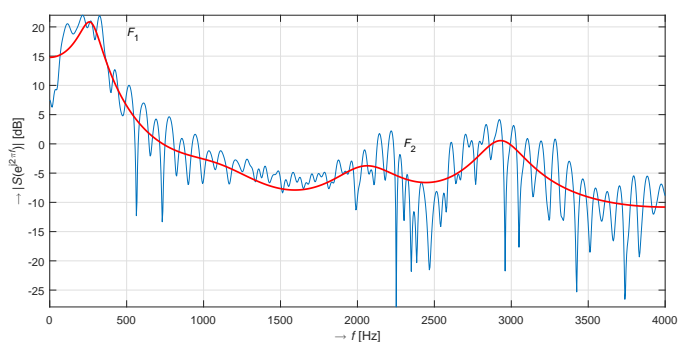
B FORMANTY ČESKÝCH SAMOHLÁSEK



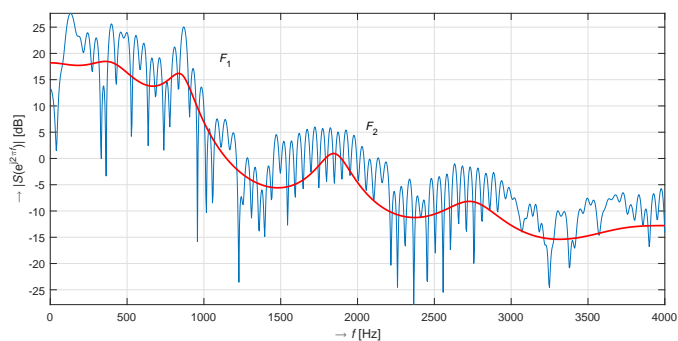
Obr. B.1: Modulové spektrum české samohlásky [a] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC. Parametry: $Fs = 8000$ Hz.



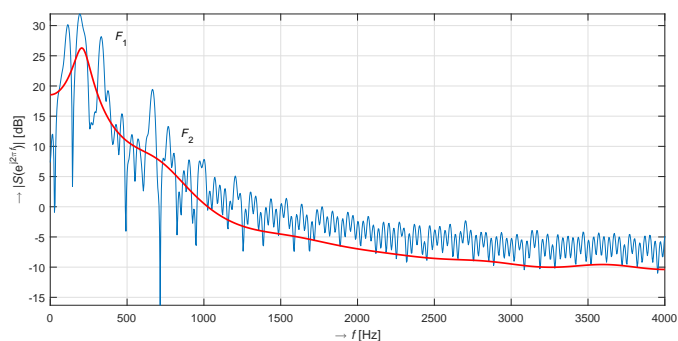
Obr. B.2: Modulové spektrum české samohlásky [e] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC. Parametry: $Fs = 8000$ Hz.



Obr. B.3: Modulové spektrum české samohlásky [i] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC. Parametry: $Fs = 8000$ Hz.



Obr. B.4: Modulové spektrum české samohlásky [o] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC. Parametry: $Fs = 8000$ Hz.



Obr. B.5: Modulové spektrum české samohlásky [u] se zvýrazněnými formantovými oblastmi proloženými funkcí LPC. Parametry: $Fs = 8000$ Hz.

C ZDROJOVÉ KÓDY PARAMETRŮ

C.1 Kód parametru funkce Formant Variability

Výpis C.1: Zdrojový kód parametru FV v programovém prostředí Matlab.

```
1 function V = formant_variability(F, fs, semitones)
2 % This function computes formant variability, which is defined as the series
3 % of absolute differences between consecutive F values (the values define
4 % F for segments) that do not equal to zero, trimmed by the median of the
5 % series of the differences.
6 % This function has the optional parameter "semitones", which can be used
7 % to convert the zero-trimmed F series to semitone scale.
8 % -----
9 % Input parameters:
10 % -----
11 % F          - input F series [Hz]
12 % fs         - sampling frequency [Hz]
13 % semitones - conversion of F to semitones (default: false)
14 % -----
15 % Output parameters:
16 % -----
17 % V          - output pitch variability structure
18
19 %% Paths and variables
20 if ((nargin < 3) || (isempty(semitones)))
21     semitones = false;
22 end
23
24 %% Trim the F0 series (by zero)
25 F = F(F ~= 0);
26 %% Convert to semitones (default: false)
27 if (semitones)
28     F = 12*log2(F(2:end)./F(1:(end - 1)));
29 end
30
31 %% Compute the absolute differences
32 D = abs(diff(F));
33
34 %% Compute the pitch variability series
35 S = D(D >= median(D));
36
37 %% Compute the output features
38 V = struct();
39
40 % Set the function series
41 V.variability_series = S;
42
43 % Set the peak features
44 [peak_values, ~, ~, ~] = findpeaks(S, fs, 'MinPeakHeight', 1.5*iqr(S));
45
46 V.num_peaks = length(peak_values);
47
48 % Set the slope of cumulative function
49 amps = S/max(S);
50 cums = cumsum(amps)/length(amps);
51 samp = 0:(length(cums) - 1);
```

```

52
53 cums = cums(:);
54 samp = samp(:);
55
56 m_c = mean(cums);
57 m_x = mean(samp);
58
59 V.cumsum_slope = sum((samp - m_x).*(cums - m_c))/sum((samp - m_x).^2);
60 end

```

C.2 Kód parametru funkce Speech Articulation

Výpis C.2: Zdrojový kód parametru SA v programovém prostředí Matlab.

```

1 function A = speech_articulation(y, fs, time_step, max_frm, win_len, preem)
2 % This function computes "articulation function" and "articulation peaks".
3 % Articulation function is defined as the product of absolute differences
4 % of the formant components computed for intervals of selected length
5 % (default: 0.5 seconds).
6 % The formant components are computed as the median formant divided by the
7 % maximum formant value per segment. The articulation function therefore
8 % describes how formants do change in time. The peaks of the function then
9 % determines the number of significant changes of articulation for a given
10 % threshold (default thr. is set to 1.5*mean(articulation function)).
11 % -----
12 % Input parameters:
13 % -----
14 % y          - input columns vector (speech signal)
15 % fs         - input sampling frequency [Hz]
16 % time_step  - time step in seconds (default: 0)
17 % max_frm    - maximum formant in Hz (default: 5500)
18 % win_len    - window length in s (default: 0.025)
19 % preem      - pre-emphasis from Hz (default: 50)
20 % -----
21 % Output parameters:
22 % -----
23 % A          - output articulation function peaks structure
24 %% Paths and variables
25 if ((nargin < 3) || (isempty(time_step)))
26     time_step = 0;
27 end
28 if ((nargin < 4) || (isempty(max_frm)))
29     max_frm = 5500;
30 end
31 if ((nargin < 5) || (isempty(win_len)))
32     win_len = 0.025;
33 end
34 if ((nargin < 6) || (isempty(preem)))
35     preem = 50;
36 end
37 %% Segment the signal
38 seg_window = hamming(0.5*fs);
39 seg_overlap = floor(length(seg_window)/2);
40 seg_padding = 0;
41 Y = segmentation(y, seg_window, seg_overlap, seg_padding);
42 %% Process the segments

```



```

43 F1_components = nan(size(Y, 2), 1);
44 F2_components = nan(size(Y, 2), 1);
45 F3_components = nan(size(Y, 2), 1);
46
47 for i = 1:size(Y, 2)
48
49     % Compute the formants
50     F = praat_formants(Y(:, i), fs, time_step, max_fm, win_len, preem);
51
52     % Compute the F1 components
53     F1 = F{1,1};
54     F1 = F1(F1 ~= 0);
55
56     nans = isnan(F1);
57     infs = isinf(F1);
58
59     if ((length(F1) > 2) && (~(sum(nans) > 0)) && (~(sum(infs) > 0)))
60         F1_components(i) = median(F1(:))/max(abs(F1(:)));
61     end
62
63     % Compute the F2 components
64     F2 = F{1, 3};
65     F2 = F2(F2 ~= 0);
66
67     nans = isnan(F2);
68     infs = isinf(F2);
69
70     if ((length(F2) > 2) && (~(sum(nans) > 0)) && (~(sum(infs) > 0)))
71         F2_components(i) = median(F2(:))/max(abs(F2(:)));
72     end
73
74     % Compute the F3 components
75     F3 = F{1, 6};
76     F3 = F3(F3 ~= 0);
77
78     nans = isnan(F3);
79     infs = isinf(F3);
80
81     if ((length(F3) > 2) && (~(sum(nans) > 0)) && (~(sum(infs) > 0)))
82         F3_components(i) = median(F3(:))/max(abs(F3(:)));
83     end
84 end
85
86 %% Post-process the parameters
87 leave = (~isnan(F1_components) & ...
88         ~isnan(F2_components) & ...
89         ~isnan(F3_components));
90
91 F1_components = F1_components(leave);
92 F2_components = F2_components(leave);
93 F3_components = F3_components(leave);
94
95 %% Compute the articulation function series
96 F1_diffs = abs(diff(F1_components));
97 F2_diffs = abs(diff(F2_components));
98 F3_diffs = abs(diff(F3_components));
99
100 S = (F1_diffs.*F2_diffs.*F3_diffs)/sum(leave);
101

```

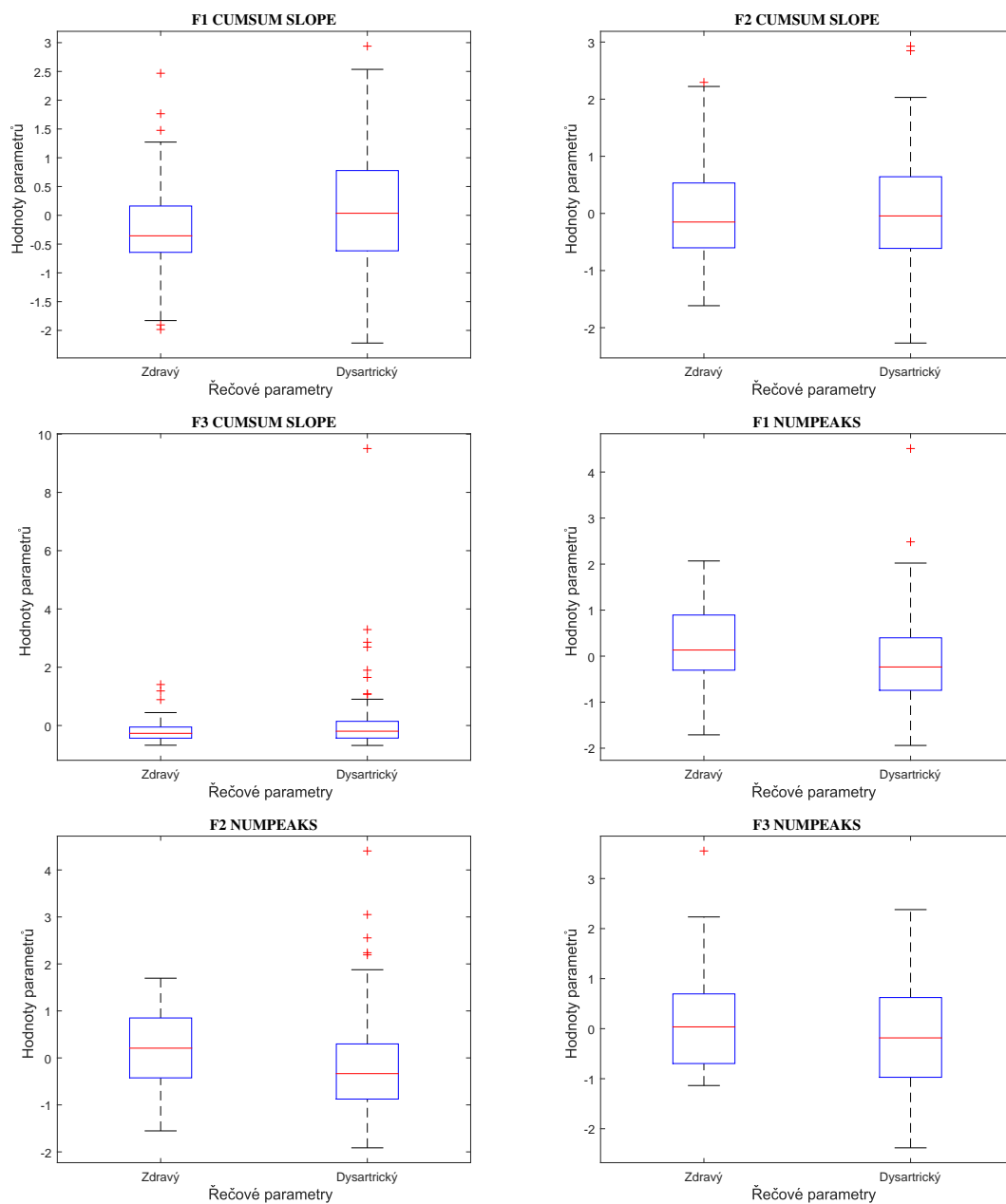
```

102 %% Compute the output features
103 A = struct();
104
105 % Set the function series
106 A.articulation_series = S;
107
108 % Set the peak features
109 [peak_values, ~, ~, ~] = findpeaks(S, fs, 'MinPeakHeight', 1.5*mean(S));
110
111 A.num_peaks = length(peak_values);
112
113 % Set the slope of cumulative function
114 amps = S/max(S);
115 cums = cumsum(amps)/length(amps);
116 samp = 0:(length(cums) - 1);
117
118 cums = cums(:);
119 samp = samp(:);
120
121 m_c = mean(cums);
122 m_x = mean(samp);
123
124 A.cumsum_slope = sum((samp - m_x).*(cums - m_c))/sum((samp - m_x).^2);
125 end

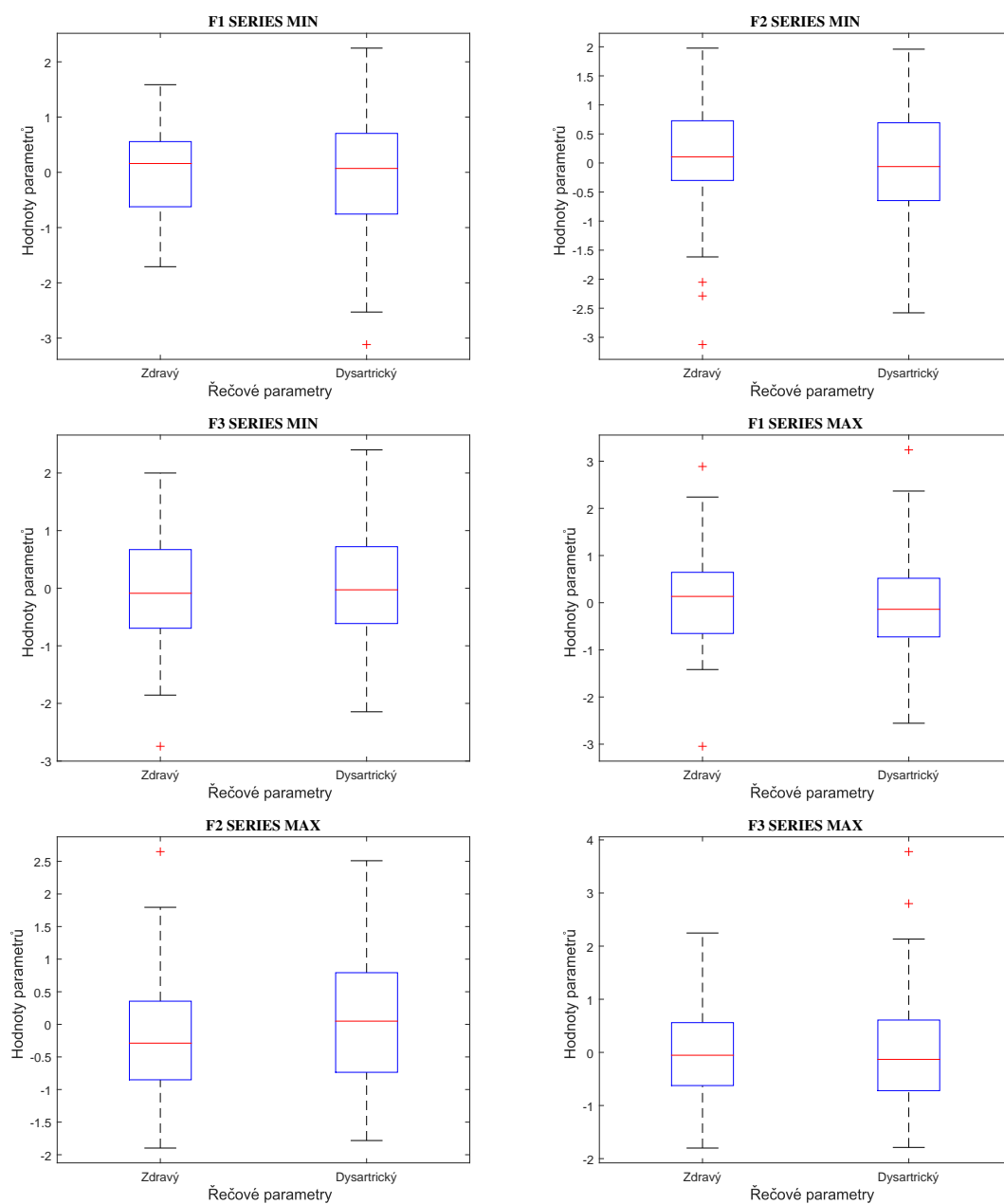
```

D POROVNÁNÍ PARAMETRŮ

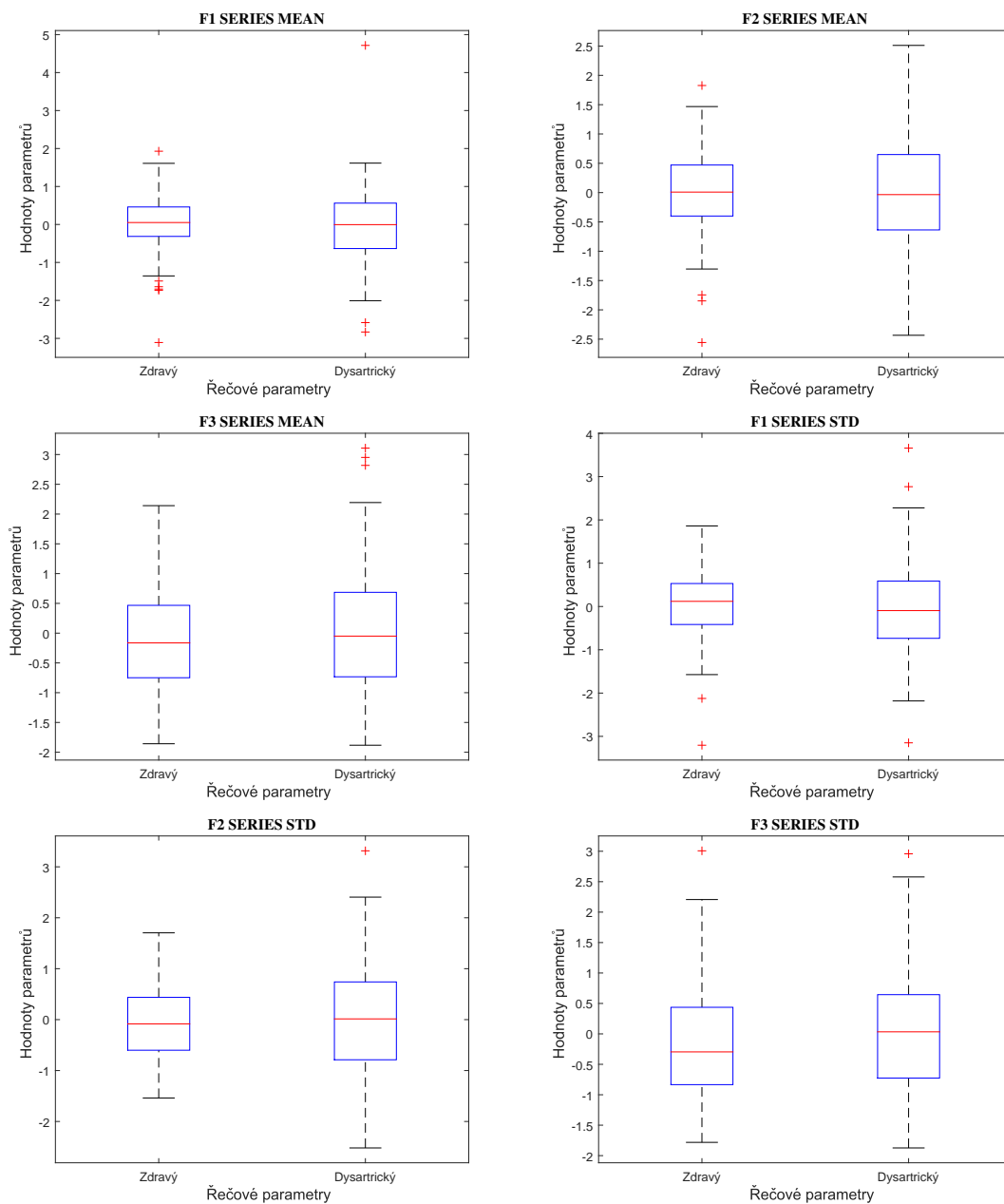
D.1 Formant variability



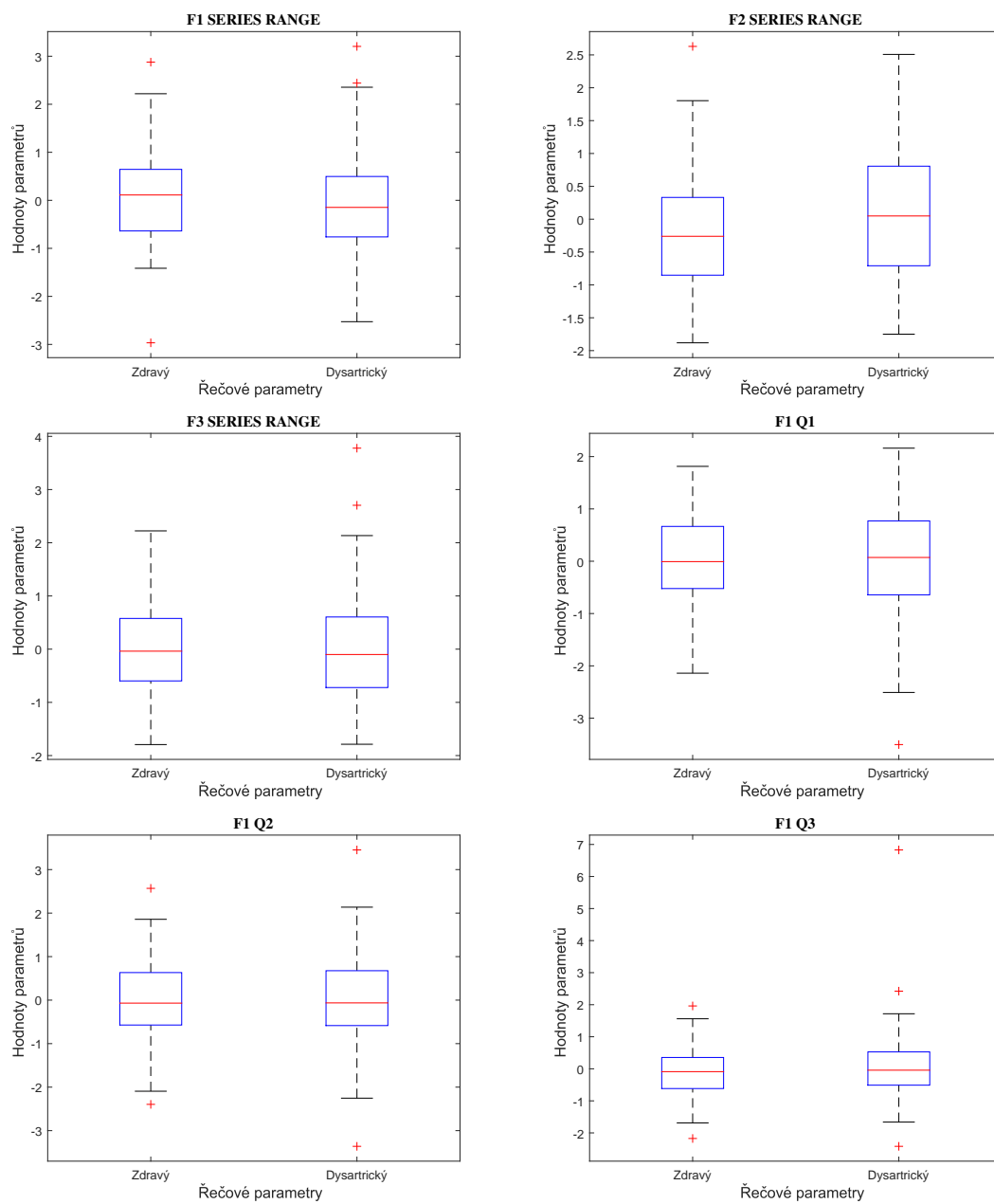
Obr. D.1: Sklon kumulativních funkcí a počet špičkových hodnot pro F_1 , F_2 , F_3 .



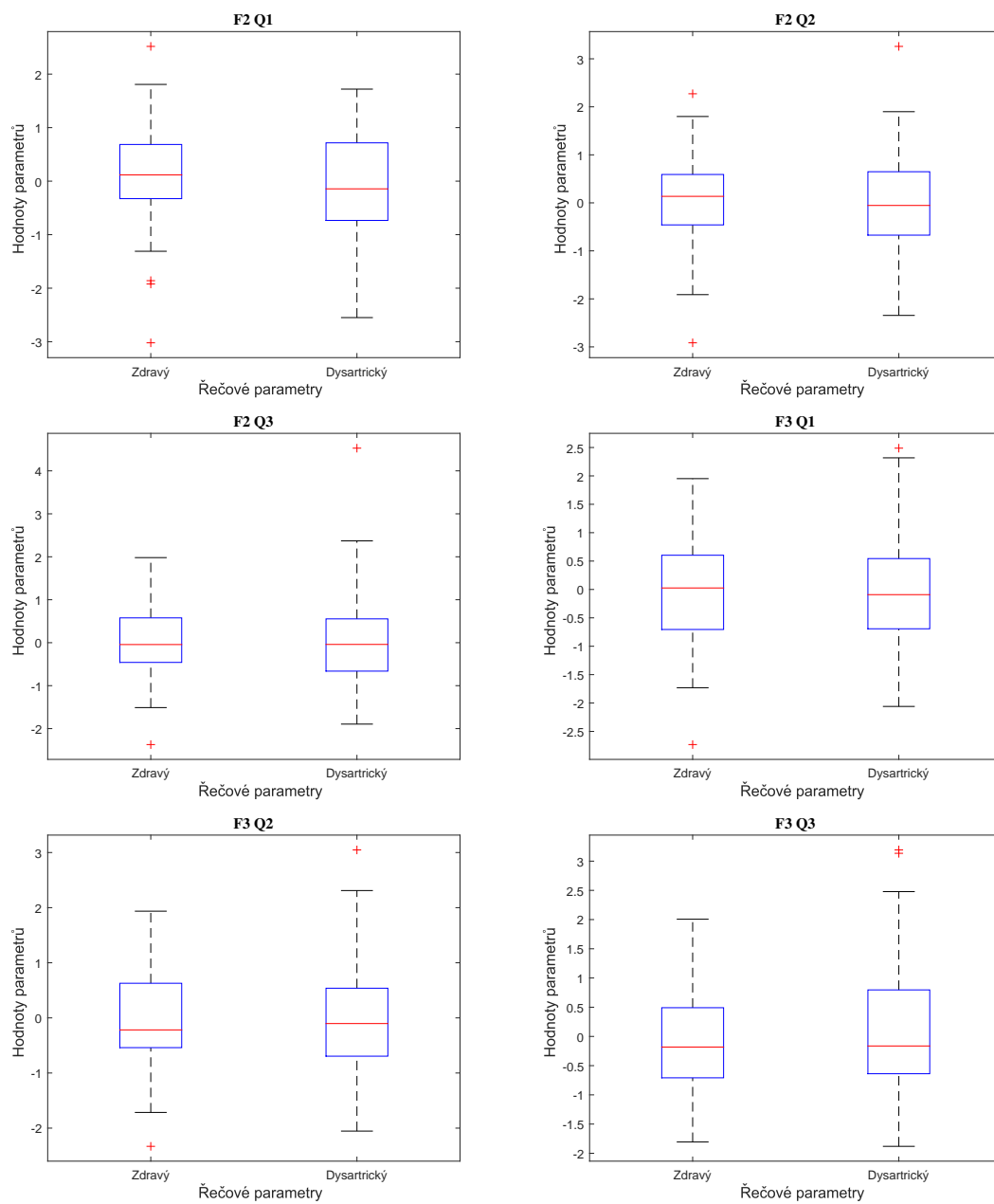
Obr. D.2: Minima a maxima pro F_1 , F_2 , F_3 .



Obr. D.3: Průměr a směrodatná odchylka pro F_1 , F_2 , F_3 .

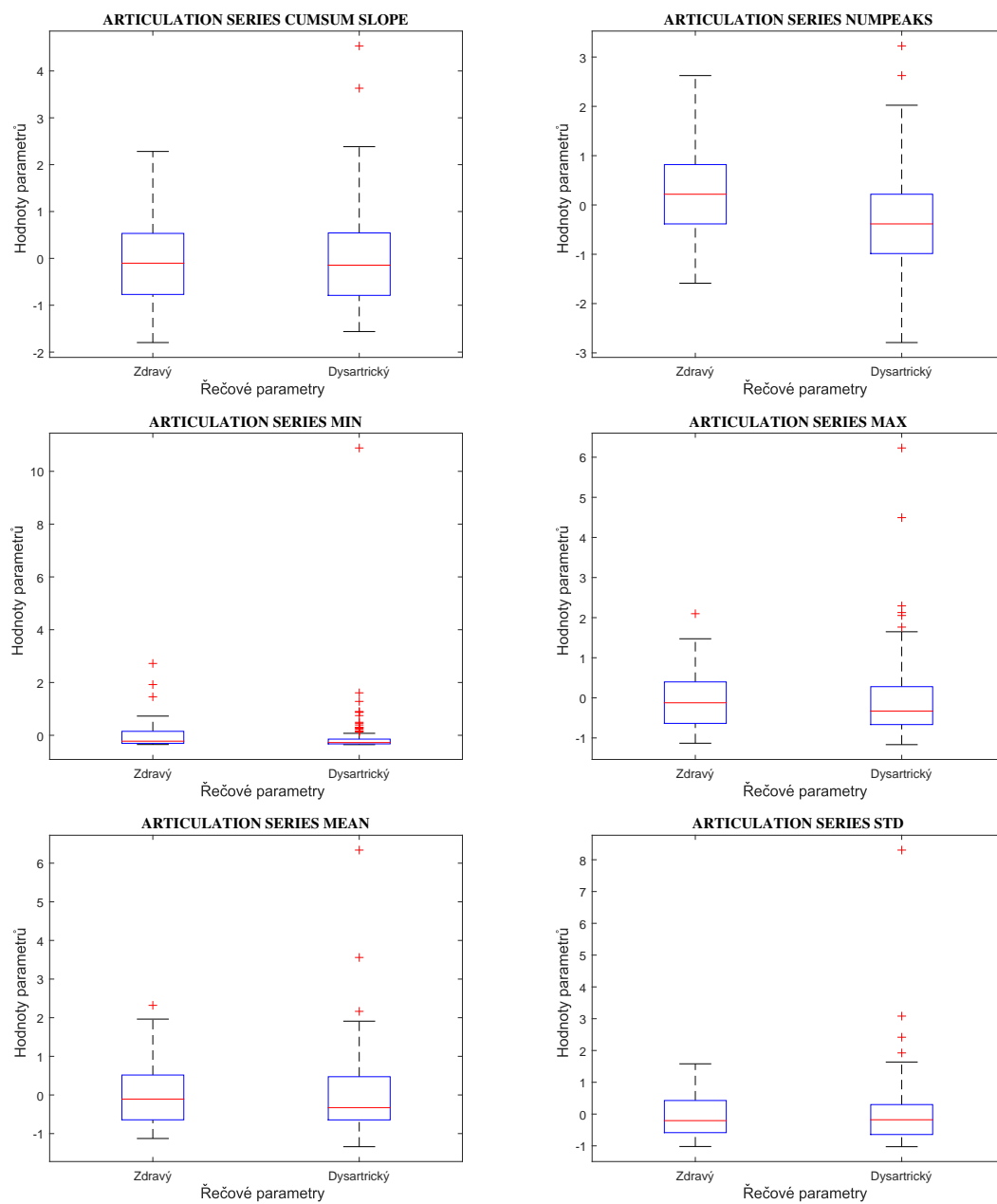


Obr. D.4: Rozsah pro F_1 , F_2 , F_3 a hodnoty kvartilů Q_1 , Q_2 a Q_3 pro F_1 .

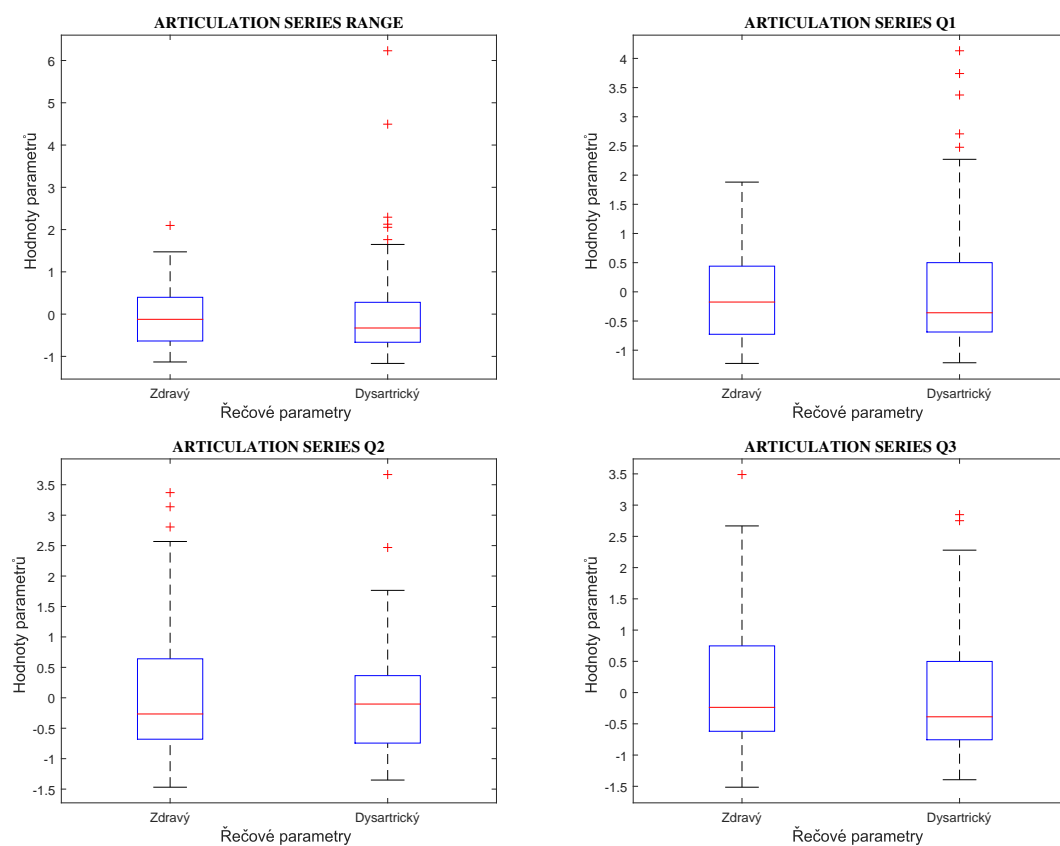


Obr. D.5: Hodnoty kvartilů Q_1 , Q_2 a Q_3 pro F_2 a F_3 .

D.2 Speech articulation



Obr. D.6: Sklon kumulativní funkce, špičkové hodnoty, minima, maxima, průměr a směrodatná odchylka pro F .



Obr. D.7: Rozsah a hodnoty Q_1 , Q_2 a Q_3 pro F .

E OBSAH PŘILOŽENÉHO CD

Na přiloženém CD se nachází:

- složka **scripts** obsahující veškeré zdrojové kódy navržené v této práci
- složka **tools** obsahující veškeré externí využité nástroje jako je software PRAAT, nástroje pro klasifikační a regresní analýzu apod.
- složka **results** s výsledky všech provedených analýz ve formě dokumentů Excel
- složka **pics** obsahující veškerou grafiku využitou v této práci
- dokument *BPVrba.pdf* obsahující text celé práce